



Designing Household Survey  
Questionnaires for Developing Countries

Edited by Margaret Grosh and Paul Glewwe

20731  
May 2000

Volume one

*Designing*  
Household Survey  
**Questionnaires** for  
Developing Countries

Lessons from 15 years of the  
*Living Standards Measurement Study*

Edited by Margaret Grosh and Paul Glewwe



The World Bank

Oxford



The  
World  
Bank



**Volume one**

# *Designing*

Household Survey

**Questionnaires** for

Developing Countries

---

Lessons from 15 years of the  
*Living Standards Measurement Study*

---

**Edited by Margaret Grosh and Paul Glewwe**



**The World Bank**

“Household surveys are essential for the analysis of most policy issues. This book has carefully assessed recent experience and developed today’s best-practice technique for household surveys. Indeed, much of *this technique* was developed and pioneered by the authors. This book is clear, systematic, and well structured. It is also wise and scholarly. It will be indispensable to anyone involved in carrying out or analyzing household surveys, and thus it is required reading for all those who wish to take evidence seriously when they think about policy.”

—Nicholas Stern, senior vice president, Development Economics  
and chief economist, the World Bank

“This book is an ambitious undertaking, but it quickly exceeded my expectations. It has many strengths: ...comprehensiveness, ...emphasis on practical application, ...and a sense of balance. For both my domestic and international survey research, this volume will serve as a valued reference tool that I will consult regularly.”

—David R. Williams, professor of sociology and senior research scientist,  
Survey Research Center, University of Michigan

“This is a comprehensive guide to planning household surveys on a range of socioeconomic topics in developing countries. It is authoritative, clear, and balanced. The work is a valuable addition to the library of any survey statistician or data analyst concerned with socioeconomic surveys in the developing world.”

—William Seltzer, former head,  
United Nations Statistical Office

Household survey data are essential for assessing the impact of development policy on the lives of the poor. Yet for many countries household survey data are incomplete, unreliable, or out of date. This handbook is a comprehensive treatise on the design of multitopic household surveys in developing countries. It draws on 15 years of experience from the World Bank’s Living Standards Measurement Study surveys and other household surveys conducted in developing countries.

The handbook covers key topics in the design of household surveys, with many suggestions for customizing surveys to local circumstances and improving data quality. Detailed draft questionnaires are provided in written and electronic format to help users customize surveys.

This handbook serves several audiences:

- **Survey planners** from national statistical and planning agencies, universities, think tanks, consulting firms and international organizations.
- Those working on either **multitopic** or **topic-specific** surveys.
- **Data users**, who will benefit from understanding the challenges, choices, and tradeoffs involved in data collection.

*Designing*

Household Survey

**Questionnaires** for

Developing Countries

---

Lessons from 15 years of the  
*Living Standards Measurement Study*

---

Edited by Margaret Grosh and Paul Glewwe

Copyright © 2000

The International Bank for Reconstruction  
and Development/THE WORLD BANK  
1818 H Street, N.W.  
Washington, D.C. 20433, U.S.A.

All rights reserved

Manufactured in the United States of America  
First printing May 2000

The findings, interpretations, and conclusions expressed in this paper are entirely those of the author(s) and should not be attributed in any manner to the World Bank, to its affiliated organizations, or to members of its Board of Executive Directors or the countries they represent. The World Bank does not guarantee the accuracy of the data included in this publication and accepts no responsibility for any consequence of their use. The boundaries, colors, denominations, and other information shown on any map in this volume do not imply on the part of the World Bank Group any judgment on the legal status of any territory or the endorsement or acceptance of such boundaries.

The material in this publication is copyrighted. The World Bank encourages dissemination of its work and will normally grant permission promptly.

Permission to photocopy items for internal or personal use, for the internal or personal use of specific clients, or for educational classroom use, is granted by the World Bank, provided that the appropriate fee is paid directly to Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, U.S.A., telephone 978-750-8400, fax 978-750-4470. Please contact the Copyright Clearance Center before photocopying items.

For permission to reprint individual articles or chapters, please fax your request with complete information to the Republication Department, Copyright Clearance Center, fax 978-750-4470.

All other queries on rights and licenses should be addressed to the World Bank at the address above or faxed to 202-522-2422.

ISBN:0-19-521595-8

**Library of Congress Cataloging-in-Publication Data has been applied for.**

# Contents

**Foreword** ix

**Acknowledgments** xi

**Contributors** xiii

## **Volume I**

### **Part 1 Survey Design**

1. Introduction 5  
*Margaret Grosh and Paul Glewwe*
2. Making Decisions on the Overall Design of the Survey 21  
*Margaret Grosh and Paul Glewwe*
3. Designing Modules and Assembling Them into Survey Questionnaires 43  
*Margaret Grosh, Paul Glewwe, and Juan Muñoz*

### **Part 2 Core Modules**

4. Metadata—Information about Each Interview and Questionnaire 77  
*Margaret Grosh and Juan Muñoz*
5. Consumption 91  
*Angus Deaton and Margaret Grosh*
6. Household Roster 135  
*Paul Glewwe*
7. Education 143  
*Paul Glewwe*
8. Health 177  
*Paul J. Gertler, Elaina Rose, and Paul Glewwe*
9. Employment 217  
*Julie Anderson Schaffner*
10. Anthropometry 251  
*Harold Alderman*

## CONTENTS

11. Transfers and Other Nonlabor Income 273  
*Andrew McKay*
12. Housing 293  
*Stephen Malpezzi*
13. Community and Price Data 315  
*Elizabeth Frankenberg*

## Volume 2

### Part 3 Additional Modules

14. Environmental Issues 5  
*Dale Whittington*
15. Fertility 31  
*Indu Bhushan and Raylynn Oliver*
16. Migration 49  
*Robert E. B. Lucas*
17. Should the Survey Measure Total Household Income? 83  
*Andrew McKay*
18. Household Enterprises 105  
*Wim P. M. Vijverberg and Donald C. Mead*
19. Agriculture 139  
*Thomas Reardon and Paul Glewwe*
20. Savings 183  
*Anjini Kochar*
21. Credit 211  
*Kinnon Scott*
22. Time Use 249  
*Andrew S. Harvey and Maria Elena Taylor*

### Part 4 Special Topics

23. Recommendations for Collecting Panel Data 275  
*Paul Glewwe and Hanan Jacoby*
24. Intrahousehold Analysis 315  
*Nobuhiko Fuwa, Shahidur R. Khandker, Andrew D. Mason, and Tara Vishwanath*
25. Qualitative Data Collection Techniques 337  
*Kimberly Chung*
26. Basic Economic Models and Econometric Tools 365  
*Jere R. Behrman and Raylynn Oliver*

## Volume 3 Draft Questionnaire Modules

- Introduction I
- Module for Chapter 4: Metadata 5  
*Margaret Grosh and Juan Muñoz*
- Module for Chapter 5: Consumption 15  
*Angus Deaton and Margaret Grosh*

Module for Chapter 6: Household Roster	31
<i>Paul Glewwe</i>	
Module for Chapter 7: Education	37
<i>Paul Glewwe</i>	
Module for Chapter 8: Health	73
<i>Paul J. Gertler, Elaina Rose, and Paul Glewwe</i>	
Module for Chapter 9: Employment	147
<i>Julie Anderson Schaffner</i>	
Module for Chapter 10: Anthropometry	219
<i>Harold Alderman</i>	
Module for Chapter 11: Transfers and Other Nonlabor Income	221
<i>Andrew McKay</i>	
Module for Chapter 12: Housing	229
<i>Stephen Malpezzi</i>	
Module for Chapter 13: Community Data	247
<i>Elizabeth Frankenberg</i>	
Module for Chapter 14: Environment	285
<i>Dale Whittington</i>	
Module for Chapter 15: Fertility	325
<i>Indu Bhushan and Raylynn Oliver</i>	
Module for Chapter 16: Migration	333
<i>Robert E. B. Lucas</i>	
Module for Chapter 18: Household Enterprise	349
<i>Wim P. M. Vijverberg and Donald C. Mead</i>	
Module for Chapter 19: Agriculture	407
<i>Thomas Reardon and Paul Glewwe</i>	
Module for Chapter 20: Savings	453
<i>Anjini Kochar</i>	
Module for Chapter 21: Credit	461
<i>Kinnon Scott</i>	
Module for Chapter 22: Time Use	483
<i>Andrew S. Harvey and Maria Elena Taylor</i>	
Module for Chapter 23: Panel Data	495
<i>Paul Glewwe and Hanan Jacoby</i>	



## Foreword

Multitopic household surveys have become an indispensable instrument for understanding development. They are fundamental to serious microeconomic analysis of the incentive and distributional aspects of policy, and therefore to the analysis of most policy issues. Researchers draw on them to test behavioral theories. Policymakers need them to assess public interventions. The development community uses them to locate the poor. Developing countries, without adequate household survey data, are forced to make policy decisions in an environment with many blind spots, where crucial information can be seen only dimly or not at all.

Household surveys are also expensive, both in terms of money and institutional capacity. Ultimately their value depends on their design and execution. Errors in their design or execution are wasteful, and can lead to policies that are harmful to the poor. It is therefore important to design and implement surveys correctly from the outset.

Margaret Grosh and Paul Glewwe have put together one of the most comprehensive and informative documents ever written on the design, implementation, and use of household surveys in developing countries. If you are engaged in any of these tasks, this book is essential reading.

The household surveys treated in this book truly are multitopic surveys, covering such topics as household size and composition, education, health, anthropometry, fertility, income and consumption, employment, agricultural production, household enterprises, transfers and nonlabor income, savings and credit, housing, the environment, migration, and time use. The editors have greatly increased the value of the basic approach by incorporating chapters on community data, panel data, and the allocation of resources within the household.

As the World Bank and other development organizations increase their efforts to reduce poverty and raise living standards in developing countries in the 21st century, the need for comprehensive, reliable and up-to-date information on economic and social conditions in these countries will be greater than ever. The vast store of knowledge in this book will contribute significantly to meeting this need. Failure to use this knowledge will consign policymakers to making their decisions without adequate information for many years to come, while systematic use of this knowledge will do much more for the poor than the innumerable speeches made and summits convened on their behalf.

Lyn Squire  
Director, Global Development Network  
World Bank



## Acknowledgments

A project of this size and scope depends on many people playing many roles. Space limitations preclude us from naming all of the hundreds of people who made contributions along the way, but we would like to acknowledge some of the most important.

The authors of the individual chapters deserve thanks for their gracious willingness to go through many rounds of revisions, spread over a longer time than anyone originally envisioned. Producing a book on the design of multitopic questionnaires requires much more cooperation among authors and several more iterations than does the standard edited volume. We are extremely grateful for the forbearance of the authors in this difficult process. The authors themselves were helped by a large number of peer reviewers. They are recognized in the individual chapters, but we would like to extend our thanks to them here as well.

Much of the work in these volumes was based on past practice in LSMS and other household surveys including, but not limited to, the World Fertility/Demographic and Health Surveys, the RAND Family Life surveys, the Social Dimensions of Adjustment surveys, and several special topic surveys such as household budget surveys, water and sanitation surveys, housing surveys, and time use surveys. While the authors pulled together the lessons from past experience, it is also important to acknowledge the irreplaceable contributions made by the thousands and thousands of household members who served as respondents, the dozens of agencies that implemented

the surveys, the many agencies that provided technical assistance and funding, and the academic participants who provided advice and criticism over the years.

The project as a whole was strongly supported from original vision to final printing by our immediate manager for most of that time, Emmanuel Jimenez, who provided us with useful technical input and a great deal of enthusiasm, patience, and bureaucratic support. We also greatly appreciate the support of his directors, Lyn Squire and Paul Collier. The project was primarily financed by a grant from the World Bank Research Committee (679-61), managed by Greg Ingram and administered by Clara Else.

Many people reviewed the book and project as a whole. We greatly appreciate these contributions by Pat Anderson, Jere Behrman, Elisa Lustosa Caillaux, Courtney Harold, John Hoddinott, Anna Ivanova, Alberto Martini, Raylynn Oliver, Prem Sangraula, Salman Zaidi, and three anonymous reviewers. To have input on the project as a whole from these outsiders was very helpful. In addition, participants at three workshops held at the World Bank, plus various training events sponsored jointly by the World Bank and the Inter-American Development Bank, critiqued the project while it was in progress.

In the course of creating the book, Diane Steele answered questions from all authors on the details of LSMS data sets. Fiona Mackintosh edited early drafts and helped to transform the disparate chapters into a single

## ACKNOWLEDGMENTS

whole. Lyn Tsoflias provided us with valuable research assistance. Word processing and conference logistics were ably handled by Thomas Hastings, Patricia Sader, Jim Schafer, and Daniel O'Connell. Questionnaire layout was mastered by Thomas Hastings, Andrea Ramirez, and Heidi Van Schooten. Contracting support from Liliana Longo, Selina Khan, and Patricia Sader was timely and organized. The final editing, layout, and design were handled dextrously by Meta de Coquereaumont, Wendy Guyette, Kate Hull, Daphne Levitas, Heidi Manley, Laurel Morais, and Derek Thurber, all with

Communications Development Inc. Communication with the World Bank's Publications Committee and with the publishers and printers was efficiently handled by Paola Scalabrin and Randi Park.

Finally, effusive and endless thanks to our families and friends who put up with the excessively long hours that we spent on this project, who cheered and calmed us through the frustrating times, and who helped us to bring this long project to a successful conclusion without completely losing track of other important aspects of our personal and professional lives.

## Contributors

Harold Alderman	World Bank
Jere R. Behrman	University of Pennsylvania
Indu Bhushan	Asian Development Bank
Kimberly Chung	Michigan State University
Angus Deaton	Princeton University
Elizabeth Frankenberg	RAND
Nobuhiko Fuwa	World Bank
Paul J. Gertler	University of California, Berkeley
Paul Glewwe	World Bank and University of Minnesota
Margaret Grosh	World Bank
Andrew S. Harvey	St. Mary's University, Halifax, N.S. Canada
Hanan Jacoby	World Bank
Shahidur R. Khandker	World Bank
Anjini Kochar	Stanford University
Robert E.B. Lucas	Boston University
Stephen Malpezzi	University of Wisconsin
Andrew D. Mason	World Bank
Andrew McKay	University of Nottingham, United Kingdom
Donald C. Mead	Michigan State University
Juan Muñoz	Sistemas Integrales
Raylynn Oliver	Consultant, World Bank
Thomas Reardon	Michigan State University
Elaina Rose	University of Washington
Julie Anderson Schaffner	Fletcher School of Law and Diplomacy, Tufts University
Kinnon Scott	World Bank
Maria Elena Taylor	St. Mary's University, Halifax, N.S. Canada
Wim P.M. Vijverberg	University of Texas at Dallas
Tara Vishwanath	World Bank
Dale Whittington	University of North Carolina at Chapel Hill



# Volume I



# Part I

## Survey Design



# 1

## Introduction

*Margaret Grosh and Paul Glewwe*

Accurate, up-to-date, and relevant data from household surveys are essential for governments to make sound economic and social policy decisions. Governments need these data to measure and monitor poverty, employment and unemployment, school enrollment, health and nutritional status, housing conditions, and other dimensions of living standards. They need the data to determine whether schools, health clinics, agriculture extension services, roads, electric power, and other basic services are reaching the poor and other disadvantaged groups. And analysts need household survey data to model economic behavior and thus provide answers to such important policy questions as: How would changes in food subsidies affect the population's nutritional status? Would increasing fees for public schools reduce school enrollment, and how much revenue would be raised by such fee increases? Who would participate in a new labor-intensive public works program, and what would be the net benefit for participants? How would changes in the price of fertilizer affect farmers' production of different crops?

One way to collect the data needed to answer these questions is to conduct separate household surveys on each topic—that is, to conduct a labor force (employment) survey, a health survey, a housing survey, and so forth. Alternatively, data on many different topics can be collected in a single survey. Such a “multi-topic” household survey, which has many advantages, is the type of survey considered in this book.

Household surveys are not a new invention. Stigler (1954) points out that systematic collection of data from households began over 200 years ago. The first known efforts were the collection of family budgets in England by Davies (1795) and Eden (1797). In the 1800s similar data were collected in Saxony, Prussia, Belgium, the United States, and undoubtedly other places as well. The motivation for much of this research was to focus public attention on the plight of

the poor. By the mid-1800s, generalizations about household behavior were being drawn from these data. For example, Ducpetiaux's 1855 study of 200 Belgian households was used by Ernst Engel to derive his classic law that the fraction of a household's budget devoted to food falls as income rises.

The statistical theory that supports modern survey methods was developed in the 1920s. This led to the establishment of high-caliber nationwide surveys in many countries, especially after World War II. Developing countries also participated in this phenomenon; for example, India's annual National Sample Survey began in 1950. With the advent of modern computing, and especially the appearance of powerful personal computers, the collection and analysis of household survey data has expanded rapidly in both developed and developing countries. (See Deaton

1997 for a brief review of household surveys in the 20<sup>th</sup> century.)

Since 1970 several major international programs have been organized to support the collection of household survey data in developing countries. Among the largest such programs have been the United Nations Household Survey Capability Program, the World Fertility Surveys (which later became the Demographic and Health Surveys), and the World Bank's Living Standards Measurement Study (LSMS) survey program. Other organizations, including the International Food Policy Research Institute, the RAND Corporation, and Cornell University, have also carried out household surveys in developing countries. Some U.N. organizations regularly participate in single-topic household surveys in developing countries, such as employment surveys done in collaboration with the International Labour Office. And two regional survey programs have been strongly influenced by, and indeed have grown directly out of, the World Bank's LSMS program. The first of these, the Social Dimensions of Adjustment (SDA) program for Sub-Saharan Africa, was supported by a consortium of agencies and administered by the World Bank. The second and more recent regional survey program, the Improving Surveys of Living Conditions program for Latin America, is sponsored jointly by the Inter-American Development Bank, the World Bank, and the Economic Commission for Latin America. (The Spanish name for this program is *Mejoramiento de las Encuestas de Condiciones de Vida*; it is often referred to by its Spanish acronym, MECOVI.) The surveys done under the LSMS, SDA, and MECOVI programs are all multi-topic surveys.

Because of these and other efforts, household survey data are now much more widely available than they were 10 or 20 years ago. World Bank statistics on the extent of poverty in 1985 were based on data from only 22 of 86 developing countries. Although these 22 countries accounted for 76 percent of the population of the 86 countries (Ravallion, Datt, and van de Walle 1991), it is significant that at that time no reliable data existed for three-fourths of the developing countries. Similar calculations currently underway are based on data from about 70 of 100 developing and transition countries, covering about 88 percent of the total population of the countries. Data for more than one point in time are now available for 50 countries. Coverage has grown the most in the region where it was lowest,

Sub-Saharan Africa. In the 1985 calculation only 6 percent of Sub-Saharan Africa's population was represented, while recent estimates cover 66 percent of this population (Ravallion and Chen 1998). Finally, the time lag between collection and dissemination of the data is getting smaller. In the 1985 World Development Report the average lag was 11 years, so the average survey date was 1974. Now the lag is only five years (Ravallion and Chen 1997).

The surge in the collection of household survey data in developing countries has greatly increased the demand for knowledge on how best to design and implement such surveys. Moreover, the growing number of surveys provides a vast amount of experience from which to draw lessons. Yet until now it has often been difficult for those planning a new survey—especially one in a developing country—to find out about the experiences of previous household surveys: what was tried, the factors that influenced decisionmaking, what worked, and, most importantly, what did not work. The formal literature is scattered across disciplines—statistics, economics, sociology, psychology—and often contained in conference proceedings or government document series that are not widely indexed and are seldom available outside the country where they were written. An additional limitation is that a considerable amount of the formal literature pertains to surveys in industrialized countries. While much can be learned from such literature, it is still unclear how well the literature applies to settings with lower literacy rates, different income levels and employment and consumption patterns, and differing factors that affect the social interaction of the interview.

Much of the experience of surveying in developing countries is poorly documented. Statistical institutes in developing countries have little money or staff to devote to experimentation or research; their mandate is production and their resources are few. Articles published in the formal academic literature that use the data from these surveys typically provide only a brief description of the data used. They may contain some hints about whether the data collection methods worked, but almost by definition, data collection efforts that failed usually do not lead to academic publications. Household survey questionnaires and their associated statistical abstracts (reports) contain both implicit and explicit information, but they are sometimes available only in the country in which they were

administered. Moreover, statistical abstracts tend to minimize any problems that may have been associated with a survey because the statistical agencies that produce these abstracts do not want to publicize a survey's shortcomings. In principle, the most useful information for the designers of future surveys would be the internal memoranda and informal notes of the agencies and people involved in designing and implementing past surveys. However, these are rarely filed and seldom preserved after a survey is completed, much less systematically made available to people outside the agency.

### **The Objective and Audience for this Book**

The objective of this book is to provide detailed advice on how to design multi-topic household surveys, based on the experience of past household surveys. This book will help individuals and organizations that are planning a comprehensive, multi-topic survey to define the objectives of their survey, identify the data needed to analyze those objectives, and draft questionnaires that will collect such data. These tasks are not easy, because designing such a survey for a given country (or an area within a country) usually involves a host of tradeoffs among different objectives. This book aims to help survey designers evaluate these tradeoffs, set realistic objectives, and design a survey that best fulfills those objectives.

This book was written with several target audiences in mind. The primary audience consists of the people most likely to carry out household surveys similar to the ones discussed in the book—the staff of the national statistical agencies and planning agencies responsible for their countries' household surveys. A second audience consists of individuals or groups in consulting firms or international aid agencies that advise governments on the design of household surveys. A third audience is composed of researchers or research agencies that plan to field a survey to pursue their own research objectives. A fourth audience consists of individuals or groups working on a survey intended to evaluate or monitor the impact of a development project in a particular country—either a nationwide project or a project limited to a small part of the country. A fifth audience is composed of people working on a single-topic survey, because the book can provide them with guidance on how to collect “background” information from households—

information on, for example, a household's composition, basic characteristics, and level of welfare. Finally, this book will assist researchers who use household survey data produced by others, because it will help them understand the challenges, possible options, and tradeoffs involved in data collection. Such an understanding will allow these researchers to interpret household survey data more accurately and use these data more fully.

The recommendations in this book apply to a broad range of multi-topic household surveys, reflecting the authors' expectation that future surveys in developing countries will be increasingly diverse in their purposes and content. The book provides survey designers with a wide range of options from which they can pick and choose according to both the purpose of their survey and the prevailing circumstances in the country studied. Future household surveys will, and should, evolve in ways that are hard to foresee. Thus this book should be regarded as a starting point for planning new surveys rather than as an exhaustive treatise on the way to design all future household surveys.

This book assumes that the survey designer has already decided to implement a multi-topic household survey, as opposed to a census, a qualitative study, or a single-purpose survey. Nevertheless, several chapters in this book compare the advantages and disadvantages of different data sources for studying certain topics. In addition, Chapter 25 provides a thorough discussion of qualitative data collection methods.

### **The Experience on Which This Book Is Based**

Much of this book is based on the experience of the World Bank's Living Standards Measurement Study (LSMS) program (Box 1.1), one of several recent international efforts to expand the pool of data on poverty and living standards in developing countries. The World Bank established the LSMS program in 1980 to explore ways of increasing the accuracy, timeliness, and policy relevance of household survey data collected in developing countries. Because the first LSMS surveys were designed by the World Bank for research purposes, there was little variation in these surveys' design and implementation. However, by the late 1990s LSMS surveys had been carried out in a wide range of low- and middle-income countries, with the involvement of many different national agencies and international organizations. Over time LSMS

**Box 1.1 An Introduction to LSMS Surveys**

The overall objective of LSMS surveys is to measure and study the determinants of living standards in developing countries, especially the living standards of the poor. To accomplish this objective, LSMS surveys must collect data on many aspects of living standards, on the choices that households make, and on the economic and social environment in which household members live. Much of the analysis undertaken using LSMS surveys attempts to investigate the determinants of living standards—which requires more sophisticated analytical methods than simple descriptive tables.

LSMS surveys have several characteristics that distinguish them from other surveys. One of the most important is that they use several questionnaires to collect information about many different aspects of household welfare and behavior. These consist of a household questionnaire, a community questionnaire, a price questionnaire, and, in some cases, a facilities questionnaire. (For more details on the questionnaires see Box 1.4.)

Another characteristic of LSMS surveys is that they typically have nationally representative, but relatively small, samples—usually between 2000 and 5000 households. This will yield fairly accurate descriptive statistics for the country as a whole and for large subareas (such as rural and urban areas or a few agroclimatic zones), but usually not for political jurisdictions (such as states or provinces). The surveys' sample sizes are generally adequate for the regression methods often used for policy analysis of LSMS survey data.

Because of the complexity of most LSMS surveys, these surveys have rigorous quality control procedures to ensure that the data they gather are of high quality. These proce-

dures, which are generally difficult to implement on larger samples, usually include several key elements. Both the survey's fieldwork and its data entry are decentralized, and the people who carry out these tasks are strictly supervised. Interviewers receive extensive training (usually for about four weeks) prior to the survey. In the field, information is gathered not by asking one person all the questions about the household and its members but through a series of "mini-interviews," with each adult responding for himself or herself. This procedure minimizes any errors caused by respondent fatigue or by the use of proxy respondents. The interviewers make multiple visits to households to find any members who were not home during the interviewer's earlier visits—which also reduces the need to use proxy respondents.

There is one supervisor for every two or three interviewers. The supervisors must revisit a significant percentage (often 25 percent) of the sampled households to check on the accuracy of the interviewer's data. They must directly observe some interviews, and they must review each questionnaire in detail. Supervisors' performance of these procedures is documented, and the supervisors are in turn supervised by staff from the central office of the statistical agency. Data entry and editing are done as soon as each interview is over, either in the local field office or by a data entry operator who travels to households with the team of interviewers. As data are entered into the computer, a data entry program carries out a large number of quality checks to detect responses that are out of range or inconsistent with the other data from the questionnaire. Any problems this program detects can be verified or corrected in a subsequent visit to the household by the interviewer.

surveys have become increasingly customized to fit specific country circumstances, including policy issues, social and economic characteristics, and local household survey traditions. Each survey has also inevitably reflected the interests (and prejudices) of the individuals planning it.

The LSMS program has had its share of successes. Most importantly it has shown the feasibility of collecting comprehensive household survey data in developing countries. Since the first LSMS survey in 1985, LSMS surveys have been implemented in about 30 developing countries (Table 1.1). In some of these countries the original LSMS survey prototype was implemented in its entirety. In other countries this prototype was significantly altered to suit local circumstances. In still other countries it was used as a guide to redesign surveys that already existed. LSMS surveys were also the starting point for SDA surveys, which

have been implemented in about 20 Sub-Saharan African countries, and for the MECOVI program now in progress in eight Latin American countries.

The increase in the number of LSMS surveys and other household surveys has substantially expanded the stock of data that can be used to study poverty and, more broadly, economic and social development in developing countries. In every country where an LSMS survey has been done, the data have been used to measure and analyze poverty by the government, an international development agency, or both working together. In several countries LSMS data have directly influenced specific government policy decisions (see Box 1.2). Data from LSMS and similar surveys have also been used in hundreds of studies of developing countries, helping to extend what is known about poverty, household decisionmaking, and the impact of economic and social policy changes on household

Table 1.1 LSMS Surveys

Country	Year of first survey	Has the survey been repeated, or will it be repeated?	Number of households in sample
Albania	1996	No	1,500
Algeria	1995	No	5,900
Armenia	1996	No	4,920
Azerbaijan	1995	No	2,016
Bolivia	1989	Yes	4,330–9,160
Brazil	1996	No	5,000
Bulgaria	1995	Yes	2,000
Cambodia	1997	Yes	6,010
China (Hebei and Liaoning only)	1995	No	800
Côte d'Ivoire	1985	Yes	1,600
Ecuador	1994	Yes	4,500
Ghana	1987/88	Yes	3,200
Guyana	1992/93	No	1,800
Jamaica	1988	Yes	2,000–4,400
Kazakhstan	1996	No	2,000
Krygyz Republic	1994	Yes	2,100
Mauritania	1988	Yes	1,600
Morocco	1991	Yes	3,360–4,800
Nepal	1996	No	3,373
Nicaragua	1993	Yes	4,454
Pakistan	1991	Yes	4,800
Panama	1997	Yes	4,945
Paraguay	1997/98	Yes	5,000
Peru	1985	Yes	1,500–3,623
Romania	1994/95	Yes	31,200
South Africa	1993	No	8,850
Tajikistan	1999	No	2,000
Tanzania—Kagera	1991	No	800
Tanzania—Human Resource Development Survey	1993	No	5,200
Tunisia	1995/96	No	3,800
Turkmenistan	1997	No	2,350
Vietnam	1992/93	Yes	4,800–6,000

Source: LSMS data bank.

welfare. Many of these studies have been presented at conferences and published in books or academic journals, and have thereby shaped thinking about these issues far beyond the countries in which the data were collected.

Despite these successes, several challenges remain for LSMS surveys and other multi-topic household surveys. First and most obviously, many developing countries still have inadequate household survey data. This is true even for some of the countries that have recently fielded new surveys, including LSMS surveys. Ideally, all governments should collect data on a regular, ongoing basis in order to monitor poverty trends over time. However, survey efforts are still sporadic in many developing countries today, and many surveys have serious deficiencies such as limited question-

naires, samples that exclude rural areas, and long delays in processing the data after completing the fieldwork.

Second, improvements are needed in the process of adapting the LSMS approach to countries that have not yet implemented LSMS-type surveys. It has been difficult for people working on a survey in one country to learn from the experience of other countries that have carried out LSMS and other multi-topic surveys. Mid-level staff in government statistical agencies know the details of why particular choices were made and know how well the choices worked, but they rarely meet with their counterparts in other countries. A small pool of World Bank staff and consultants also know many of these details and have been in contact with many of the people developing new surveys in particular countries. However, until now, they have

**Box 1.2 Using LSMS Data to Inform Government Policy Choices**

LSMS household surveys are designed to collect data that can be used to study living standards and how living standards are affected by government policies. The following examples illustrate how some governments and donor agencies have used LSMS data to help make policy choices.

In 1989 the Jamaican government was considering whether it should eliminate subsidies for basic food items and use the funds saved to expand its food stamp program. While the government was making this decision, data from the Jamaican LSMS survey became available. Analysis of these data showed that most of the benefits from general price subsidies went to nonpoor households, while most of the benefits of the food stamp program went to the poor. This information helped the government decide to remove the subsidies on basic foodstuffs and expand its food stamp program. The government then commissioned further analysis of the LSMS data to find out how many families needed help in purchasing a minimum food basket, and how much help these families needed. The government used this information to choose new eligibility thresholds and benefit levels for the food stamp program.

The Jamaican government has used its LSMS data in making many other decisions, such as whether to change kerosene subsidies and whether and how to subsidize medicines distributed through public health clinics. In addition, the government has used LSMS data to study the effects of raising user fees for public health care services. The Jamaican LSMS survey is conducted annually; the incidence of poverty is measured in each survey.

In South Africa the 1993 LSMS survey provided the first comprehensive, credible data set for the entire territory of South Africa, including the homelands. The survey was completed just before the first democratic elections were held in

the country. The data were quickly put to extensive use both by the new government and by academic researchers. The first product, an extensive statistical abstract, was followed by a poverty profile prepared jointly by the World Bank and the government's Ministry of Reconstruction and Development, then by other studies and reports. This body of work has helped to shift the national debate about poverty away from the nature and extent of poverty toward policy options for reducing poverty. For example, young women in rural areas were made eligible for public works employment programs after the data showed that these women were often needy and that they would be able to participate in such schemes since they had access to childcare. The survey data also revealed that the old age pension program was well targeted, which convinced the government not to modify that program but instead to consider reforming other programs that appeared to be less well targeted.

In 1998 the Government of the Kyrgyz Republic undertook a thorough assessment of the current and projected impact of its state pension reform. With the help of a World Bank team, the government analyzed data from the 1993 and 1996 Kyrgyz LSMS surveys to examine a range of policy alternatives. The survey data were used to show rates of participation in, contributions to, and receipts from pension programs by age cohort and by level of welfare. The data were particularly helpful to the government when it worked on setting a new level for the minimum pension, based on average earnings in the poorest quintile of the population. Forthcoming analytical work will include an assessment of household consumption and demand for utility services, and the formulation of a strategy to compensate the poorest for increases in utility prices.

shared their knowledge of past surveys mostly on an informal basis, one person or one country at a time. And since the teams that are assigned to work on each specific survey are usually small, these teams start the survey development process with detailed knowledge of some of the topics to be covered by the survey questionnaires but less detailed knowledge of others.

Third, the data gathered from some parts of LSMS survey questionnaires have been disappointing. Two particularly difficult problems are how to measure household income from agriculture and nonagricultural self-employment and how to measure savings and financial assets.

Fourth, new issues have emerged since the first LSMS surveys were implemented. The economics profession has increasingly discounted the notion of the household as a unified decisionmaking body, try-

ing instead to understand how goods, services, and power are allocated among the different members of a given household. In addition, there is growing interest in using qualitative and quantitative techniques in complementary ways, or even combining these techniques. And analysts increasingly use household survey data to address environmental issues.

**How this Book Came to Be**

In recognition of the continuing challenges for LSMS surveys, in the mid-1990s the World Bank initiated the multiyear research project that developed this book. (See Box 1.3 for a brief description of related initiatives.) The project was assigned three goals: to extend the range of policy issues that can be analyzed with LSMS data; to increase the reliability and accuracy of

**Box 1.3 Other LSMS Products**

*A manual for planning and implementing the LSMS survey.* When work began on this book about questionnaire design, work also began on a companion volume about planning and implementation: "A Manual for Planning and Implementing the Living Standards Measurement Study Survey" by Margaret Grosh and Juan Muñoz. The manual, completed in 1996, is intended for all people involved in planning and implementing an LSMS survey, including staff in planning agencies, statistical agencies, line ministries, academic institutions, and development agencies. The manual discusses such issues as sampling, fieldwork, data management, initial analysis, dissemination, and a host of planning and budgeting issues—in each case explaining the technical procedures and standards used in LSMS surveys. The manual is available in English, Spanish, and Russian.

*The LSMS data bank.* Data from LSMS surveys are now much more accessible than they were in the early years of the LSMS program in the late-1980s. The LSMS website, <http://www.worldbank.org/lsmsslsmshome.html>, contains a catalogue of the data sets that are available, the documentation for most surveys, and the data from some of the surveys. Data sets and documentation not yet available from the website are available by mail. Three factors have made it possible to increase the accessibility of LSMS data. First, a growing number of countries have adopted more open data access policies, and some have even given the World Bank permission to place their data on the LSMS website. Second, the LSMS team at the World Bank has thoroughly documented most of the surveys, whether working alone, working with managers of survey projects, or commissioning documentation work. Good documentation preserves institutional memory, lowers the cost to the Bank of disseminating data, and reduces startup costs for new users of LSMS data sets. Third, the LSMS team now has a full-time data manager, good technical support, and adequate space to stock an inventory of questionnaires, manuals for field staff, abstracts, and other useful documents from each country's survey.

*Other tools.* The LSMS program periodically produces other tools for survey planners and analysts. The best way to keep abreast of these tools is to look on the "tools for managers of new surveys" and "tools for using household survey data" pages of the LSMS website. Readers of this book may be particularly interested in a paper by Deaton and Zaidi (1999) on how to construct consumption aggregates, which complements Chapter 5 in this book, and in a recent book by Deaton (1997) on analyzing household survey data, which brings together a large amount of statistical and econometric material relevant for policy analysis.

the surveys; and to make it easier to implement LSMS surveys, either by simplifying survey design or by providing more and better instructional materials on survey design and implementation. This book contributes to the achievement of all three goals and thus addresses the four challenges facing the LSMS that were described above.

Past LSMS surveys have typically consisted of a household questionnaire, a community questionnaire, and a price questionnaire; sometimes they have also included a school or health facility questionnaire. The household and community questionnaires are each composed of separate *modules*, sections of the questionnaire that focus on different topics (Box 1.4). This book reviews each module that has typically been a part of past LSMS surveys, and offers some interesting new additions.

The author or authors of each chapter of this book were chosen according to the following criteria: extensive research experience on the topic in question; experience in analyzing data on that topic using data from LSMS and non-LSMS surveys (both multi-topic and single-topic); and experience in collecting data in developing countries. In order to ensure that experiences and perspectives from both LSMS and non-LSMS surveys were included, a concerted effort was made to include not only people who have long been associated with LSMS surveys but also people associated with other survey traditions.

The authors of the chapters that focus on specific modules have reviewed the relevant literature (both analytical literature and literature on survey experience), analyzed existing survey data, and, in the case of the consumption module, experimented with different methods of collecting data. Many authors have drawn lessons not only from LSMS experience but also from experience of other surveys, including the RAND Family Life surveys, the World Bank's Social Dimensions of Adjustment (SDA) surveys, and the Demographic and Health Surveys (DHS). The authors of many chapters have reviewed a large number of single-topic surveys, including ones on housing, agriculture, water and sanitation, time use, and household income and expenditure.

While this book was being written, two workshops were held that brought together all of the authors, as well as representatives from the various organizations that constitute the main audiences for this book. The participants in the first workshop were primarily data users—researchers and policy advisors. They were invit-

**Box 1.4 Components of a Typical LSMS Survey**

One distinguishing characteristic of LSMS surveys is that they are both multi-topic and multi-level: they use several questionnaires to study many different aspects of household welfare and behavior. The largest LSMS questionnaire is the household questionnaire. The LSMS household questionnaire always collects detailed information to measure household consumption, which is the best monetary indicator of household welfare (see Chapter 5 for further discussion). The household questionnaire also collects information on income; transfer income and income from wage employment are collected in almost every LSMS survey, and many LSMS surveys also collect data on income from agriculture, household enterprises, and miscellaneous sources.

LSMS household questionnaires always record information on a variety of other dimensions of welfare and on the use of social services; housing and related amenities such as water and sanitation; the level of education of adults, grade attainment and current enrollment rate of school-aged children; and vaccination histories and anthropometric (height and weight) measurements for children. A typical household questionnaire collects more information than this, in order to expand the range of living standards indicators that can be studied and to allow researchers to model the choices households make. The traditional list of modules included in a prototype LSMS survey includes: household roster; education, health, employment, migration, anthropometry, fertility, consumption, housing, agriculture, household enterprises,

miscellaneous income, and savings and credit. Some of the information (consumption, housing quality, agricultural production) is collected only at the household level, but much of it (employment, education, health) is collected at the individual level.

The *community questionnaire* gathers information on local conditions common to all households living in the same community. Many of these conditions recorded can be directly influenced by government actions. The information covered typically includes the basic characteristics (including distance from the community) of nearby schools and health facilities, the existence and condition of local infrastructure (such as roads and public transportation), sources of fuel and water, availability of electricity, means of communication, and local agricultural conditions and practices.

A separate *price questionnaire* is used to record the prevailing prices of commonly purchased items in local shops and markets. In almost all countries prices vary considerably among regions—in order to compare the welfare levels of households that live in different regions one needs information on the prices that they face when purchasing goods and services. The community and price questionnaires are discussed in Chapter 13.

Finally, in some LSMS surveys special *facility questionnaires* are used to gather detailed information on schools or health facilities. These questionnaires are discussed in Chapters 7 and 8.

ed to ensure that the book had correctly identified the research and information needs of potential data users. A larger share of the participants in the second workshop were data producers—staff from national statistical agencies and representatives of organizations that provide technical assistance or funding to national statistical agencies. They were invited to ensure that the book addressed their concerns—a requirement for any successful survey. Prior to the first workshop each chapter was reviewed by an expert in the relevant field. Before the second workshop the draft manuscript as a whole was reviewed by several experts in analyzing and producing household survey data in developing countries. After all of these people's advice had been incorporated and a polished draft produced, the book was subject to another round of (anonymous) peer review and revisions. In addition, many of the draft chapters were given to people who were in the process of advising governments or survey institutions on the design of a multi-topic household survey. This served as a limited field test for the book, and also confirmed that government plan-

ning and statistical agencies—and advisors of these agencies—had a genuine and pressing need for this book.

Nevertheless it must be recognized that because the draft modules presented in each chapter are based primarily on lessons from past surveys, few of them have been rigorously field tested in the exact form presented here. Thus extensive field testing must be done in each country implementing a new survey. Survey designers should consider this testing a vital part of their job after they have chosen a set of modules, modified these modules, and combined the modules into survey questionnaires.

Each chapter contains a “cautionary advice” box that specifies how much the draft module has been changed from its design in previous LSMS surveys, how well similar modules have worked in the past, and which parts of the modules most need to be customized to fit specific country circumstances.

This book represents a major advance on three fronts. First, the book makes it easier for those work-

ing on new household surveys to learn from the wide range of LSMS and other survey experience. Second, the book makes it much easier to customize the design of questionnaires for new multi-topic surveys. Third, the material presented in the book deals with new policy questions, presents new analytical methods to address both new and long-standing policy issues, and provides new ways to reduce or avoid measurement problems.

### **How to Use this Book**

The process of designing a comprehensive, multi-topic survey can be divided into five steps. First, survey planners must define the fundamental objectives of the survey and decide on the overall design of the survey in light of these objectives. Second, within this general framework the survey planners must choose which modules to include in the questionnaires, the objectives of each of these modules, and the approximate length of each module. Third, the planners must work out the precise design of each module, question by question, in light of the module's specific objectives and approximate length. Fourth, the modules must be integrated with each other and combined into a complete set of draft questionnaires (household, community, price, and in some cases, facility). Fifth, the draft questionnaire should be translated (if applicable) and field tested. Ideally, the five steps should be completed in chronological order. However, in practice, implementing any given step may reveal information that requires survey designers to rethink a previous step.

This book consists of three volumes. Volumes 1 and 2 contain all 26 chapters of this book. Volume 3 provides the draft questionnaires introduced by the chapters in Volumes 1 and 2. Volumes 1 and 2 are organized into four parts. The first three chapters of Volume 1 constitute Part 1, which discusses the “big picture.” This includes decisions that must be made about the overall design of the survey and the modules to be used, as well as procedures for combining modules into questionnaires and questionnaires into a survey (or sequence of surveys). Chapter 2 starts by describing how to choose from among the three “classic” survey designs and how to select the modules to be included in the survey. Chapter 3 describes general procedures for designing each module, combining the modules into a well-integrated set of questionnaires, and translating and field testing the questionnaires.

The remaining chapters of Volume 1 form Part 2 of the book, and the first nine chapters of Volume 2 comprise Part 3. The chapters in Parts 2 and 3 discuss, in great detail, the individual modules that are the building blocks of any multi-topic household survey. Each chapter reviews the main policy issues pertinent to the subject matter of the module, identifies the data needed to analyze these issues, introduces one or more draft modules (which are presented in Volume 3), and provides annotated notes that explain the reasoning behind many of the details of each draft module. For most modules, two or three different versions are introduced, each of different length. Which module to use depends on the level of interest in the particular topic. In addition, many of the chapters in Parts 2 and III discuss how to add or delete submodules within each module in order to provide a better fit with local circumstances and the specific focus of the survey.

Part 2 (Chapters 4–13), in Volume 1, introduces “core” modules that must be included in virtually all LSMS-type surveys: metadata, consumption, roster, education, health, employment, anthropometry, transfers and other nonlabor income, housing, and the community questionnaire. The modules on health and education come with draft questionnaires that can be used for gathering data from local schools and health care facilities. The collection of community-level data, including data on local prices, is discussed in Chapter 13, the final chapter of Part 2.

Part 3 (Chapters 14–22), in Volume 2, introduces modules that are optional: environmental issues, fertility, migration, total income, household enterprises, agriculture, savings, credit, and time use.

The last four chapters of Volume 2 constitutes Part 4 of the book. These chapters contain material that is more methodological in nature. Chapter 23 discusses when and how to collect panel data—in other words, whether to interview the same households when doing a sequence of surveys, and how best to do so when this option is chosen. Chapter 24 reviews the issues involved in analyzing the allocation of resources and power within households, and summarizes the implications of this analysis for data collection. Chapter 25 summarizes how qualitative research methods can be used to complement the quantitative methods typically used in the design and analysis of multi-topic household surveys. This chapter stresses that qualitative methods can play a useful role in the design of multi-topic household surveys, especially in

formulating questions and developing hypotheses for data analysis. It is unfortunate that these methods have been neglected by most survey designers, who usually have quantitative backgrounds.

Chapter 26 reviews the basic economic and econometric concepts that underpin many of the chapters in this book. Although many survey designers have an economics background, many others do not, and even some economists may benefit from a review of this material. The chapter begins by presenting the basic economic model of the household and goes on to discuss standard econometric techniques that have often been used in policy research on developing countries.

The reader should understand that the questionnaires provided in Volume 3 and discussed in Parts 2 and III are not polished or completed, and cannot be used immediately in any developing country. Final versions of the questionnaires for any country must be developed by the survey designers themselves. Survey designers must combine their own experience and expertise with the information in this book to design a country-specific questionnaire that will elicit the information needed to answer the most important policy questions of that country. This book is just a starting point. It provides survey designers with the lessons learned from past experience and with advice from experts who are familiar with both LSMS and other household surveys.

### *Part 1*

All readers of this book should read all of Part 1, which in addition to this chapter includes Chapters 2 and 3.

**CHAPTER 2: MAKING DECISIONS ON THE OVERALL DESIGN OF THE SURVEY.** This chapter leads survey designers through the factors that need to be considered when determining the basic scope of the survey. It sketches three alternative designs for an LSMS-type survey—a full LSMS survey, a “core” (scaled-down) LSMS survey, and a core and rotating module design—and suggests rules to help survey designers choose the design most appropriate for the circumstances that they face. Chapter 2 also explicitly defines the “core” components that should be included in any LSMS-type survey.

**CHAPTER 3: DESIGNING QUESTIONNAIRE MODULES AND ASSEMBLING THEM INTO SURVEY QUESTIONNAIRES.** This chapter moves to a finer level of detail, providing gen-

eral guidance on how to design individual modules and combine them into an integrated set of survey questionnaires. First each module must be customized to meet the specific objectives set out for it. Then the modules must be compared with each other to check for gaps and overlaps and to harmonize wording, codes, and recall periods. Next, survey designers must decide on the order of the modules and combine the modules into draft questionnaires. Finally, the questionnaires must be translated and field tested. Throughout this process, issues of questionnaire formatting will arise; Chapter 3 explains the principles and conventions used in formatting LSMS questionnaires.

### *Part 2*

After survey designers have read Part 1 of this book and decided on the broad issues concerning the scope and design of the survey, they can begin the painstaking but crucially important task of designing the individual modules. Detailed advice on designing individual modules is given in Parts 2 and 3 of the book. Part 2 includes all modules that should be included in almost any LSMS-type survey, while the modules in Part 3 are optional. Almost all of the chapters in Parts 2 and 3 follow a similar outline. The first section reviews the current policy issues in developing countries for the topic or sector covered by the chapter. The second section explains what kinds of data are needed from household surveys to address these policy questions and also discusses any measurement issues. The third section introduces one or more versions of a draft module (the modules themselves are presented in Volume 3), and the fourth section provides notes that explain the reasoning behind, and important details of, each version of the draft module.

Chapters 4 on metadata and 5 on consumption are the first chapters in Part 2 because they contain a good deal of information on survey methods and issues of validity and measurement, information with broad implications for the subsequent chapters. Like Chapters 4 and 5, Chapters 6–13 cover “core” modules. Each of these topics does not have to be covered in great detail, but it is recommended that at least the essential parts of each of these modules be included. For example, the fullest version of the health module introduced by Chapter 8 is so extensive that it should be included only in a survey specializing in health issues. Yet questions 10–38 of the short health module are an essential part of the core.

**CHAPTER 4: METADATA—INFORMATION ABOUT EACH INTERVIEW AND QUESTIONNAIRE.** “Metadata” are data about the survey itself, such as dates of interviews, identities of respondents, and time required for each interview. This topic has frequently been neglected in LSMS and other household surveys. Metadata are needed to guide the implementation of a survey, to help analysts interpret survey data, and to allow a quantitative evaluation of different survey procedures. Chapter 4 reviews the different kinds of metadata that should be collected as part of any multi-topic household survey, and provides guidance on how to collect them. If the recommendations made in this chapter had been adopted at the beginning of the LSMS program, the rest of this book and its companion manual (Grosh and Muñoz 1996) would have had a firmer empirical foundation for discussing the tradeoffs in survey design and implementation. For example, information on the time required to complete specific modules of varying lengths would have been very useful for deciding costs of expanding a module in terms of interview time.

**CHAPTER 5: CONSUMPTION.** This chapter differs from the others by focusing most of its attention on measurement issues, specifically on how to measure household consumption. The chapter draws on the literature on data collection, and also on data collection experiments that were part of the research for this book. One important conclusion of Chapter 5 is that measurement of consumption is highly sensitive to differences in methods, and that consumption measurement techniques within a given country must therefore be standardized over time. Researchers and consumers of household survey data need to be aware of the stringent comparability requirements that consumption data must satisfy before comparisons can be made across different surveys.

**CHAPTER 6: HOUSEHOLD ROSTER.** One of the fundamental decisions in any household survey is deciding who is and who is not a household member. Chapter 6 provides a basis for identifying all members of the household, thus selecting from which individuals the survey will collect information. The recommendations in this chapter do not differ greatly from procedures used in past LSMS surveys. The chapter offers no new proposals on how to define a “household,” which is a difficult issue in almost any household survey. LSMS

surveys have not made any significant contribution to this issue. The chapter confirms the usefulness of gathering household roster information on any children and parents of household members who do not live in the household, as well as of linking parents and children to each other when both are household members.

**CHAPTER 7: EDUCATION.** This chapter recommends only modest changes to the design of the education module that has been used in most previous LSMS surveys, because this design has worked quite well in the past. The education module (presented in Volume 3) collects information about the schooling of all household members, including highest grade attained, degrees obtained, and grades repeated. Individuals currently in school are also asked about the type of school they attend (public or private), their recent attendance, and the amount of money the household spends on their schooling. Chapter 7 also introduces an expanded version of the education module, which is useful for the design of surveys that focus on education issues. This expanded module requires administering relatively short tests of cognitive skills to members of the household as well as collecting information about local schools through school and teacher questionnaires.

**CHAPTER 8: HEALTH.** The health data collected in previous LSMS surveys have been of limited usefulness for policy analysis. The draft health module introduced in Chapter 8 (and provided in Volume 3) has been dramatically revised from previous surveys. The new module consists of a series of submodules on self-reported health status, health-related behavior, child immunization, insurance coverage, health service utilization and cost, and knowledge of health providers. One key change is that information is collected on all visits made by household members to medical facilities during the reference period—not just the most recent visit. Another improvement is the collection of self-reported data on “activities of daily living,” data which cover the ability to climb stairs, carry heavy loads, or walk long distances. The chapter also presents an expanded version of the module that includes questions related to mental health and very detailed questions on the household’s health expenditures and utilization of health facilities. The expanded version of the module also collects data on observed activities of daily living and the cognitive functioning of house-

hold members that are observed by the interviewer.

**CHAPTER 9: EMPLOYMENT.** The collection of data on employment and labor force participation in past LSMS surveys has been fairly successful. However, given the large size of this module, there is ample room for many small improvements. Several modifications are suggested to past designs. First, detailed information on household members' work for household businesses or on the household farm is collected in the household business and agriculture modules, not in the employment module. Second, job history information is collected in a way that focuses on each individual's employment five years before the time of the survey. Third, summary employment information is collected in a way that can accommodate individuals who have done many different kinds of work in the past 12 months.

**CHAPTER 10: ANTHROPOMETRY.** Nutritional status has always been one of the key nonmonetary indicators of welfare in LSMS surveys, especially for children; future surveys should continue to collect such data. Chapter 10 discusses the tradeoffs involved when anthropometric information is collected only for young children, rather than for all household members. The chapter recommends that in general anthropometric data should be collected for all household members, both adults and children. Chapter 10 also discusses the merits of collecting data on mid-upper arm circumference, which until now has been done in only one LSMS survey.

**CHAPTER 11: TRANSFERS AND OTHER NONLABOR INCOME.** Many households receive income unrelated to any of their members' current work activities. Past LSMS surveys have usually collected data on this income in a module on transfers and other nonlabor income. Chapter 11 introduces an improved version of that module, capturing more information on both public and private transfers. For private transfers, the new module also collects more information about the donor household and about the purpose of the transfers.

**CHAPTER 12: HOUSING.** Past LSMS surveys collected data on housing to serve as indicators of "basic needs" and to derive the implicit consumption value (imputed rent) associated with owner-occupied housing. Chapter 12 introduces two draft housing modules, a

short module and a longer module. The short module collects data similar to those collected by the housing module used in previous LSMS surveys, albeit with several improvements. The longer module collects information that can be used to study a wide range of housing policy issues. Both modules are flexible in that they can gather data on complex water supply systems and on many different rental arrangements. Finally, in the new housing module questions are added that are appropriate to places with cold climates and well developed infrastructure; such questions will be particularly useful for surveys in the transition economies of Eastern Europe and the former Soviet Union.

**CHAPTER 13: COMMUNITY AND PRICE DATA.** Community questionnaires have been used in many past LSMS surveys to gather information on the economic environment in which households operate. Community characteristics that affect households' economic environment can often be directly changed by government interventions. This chapter provides a much-needed discussion of how to define the "community" for which the information is to be collected and how to gather community information from a group of respondents. Finally, it introduces a longer and much more comprehensive community questionnaire (presented in Volume 3) than has been used in previous LSMS surveys. The design of this draft questionnaire is based on the experience of both LSMS surveys and the RAND Corporation's Family Life Surveys, as well as on suggestions from many of the authors of the other chapters in this book.

### *Part 3*

This part, consisting of Chapters 14–22, is in Volume 2. The chapters in Part 3 follow the same format as those in Part 2. Part 3 covers topics that are likely to be of interest, but one would never include all of them in any one survey. None of the topics in Part 3 are required for an LSMS-type survey.

**CHAPTER 14: ENVIRONMENTAL ISSUES.** To date, very few LSMS surveys have collected data that can be used to examine environmental issues. The environmental module introduced in Chapter 14 (and presented in Volume 3) offers a series of submodules that can be used, as appropriate, in different settings. There are very brief submodules on environmental priorities in urban and rural areas, on attitudes and perceptions

about urban air quality, and on discount rates. All of these submodules could be used in many surveys, even surveys that do not focus on the environment. The environmental module also includes lengthy submodules on water, sanitation, and fuel, to be included in surveys for which the use of these resources is of particular interest. There are also contingent valuation submodules that attempt to measure the extent to which households are willing to pay for improvements in urban air quality, the urban water supply, urban sanitation, or the rural water supply. The design of all of the above submodules is based on extensive experience from single-purpose surveys.

**CHAPTER 15: FERTILITY.** This chapter follows the same general approach used in many past LSMS surveys. The chapter introduces a short version of a fertility module that collects the data necessary to understand some general aspects of contraceptive use and to compile a maternity history that lists all births. Chapter 15 also introduces a standard version that includes a maternity history, a reproductive health submodule covering the previous three years, a longer section on contraceptive use, and a section on fertility preferences. Both the short and standard versions are presented in Volume 3. Whichever version of the module survey designers choose to use, the module should be administered to all women in the household of child-bearing age. This departs from the past LSMS practice of interviewing only one randomly selected woman per household.

**CHAPTER 16: MIGRATION.** Data on migration have been collected in many past LSMS surveys, but the amount of information collected has been quite small and the data have rarely been analyzed, despite significant interest in migration among researchers. This chapter introduces three different versions of the migration module: a short version, a standard version, and an expanded version. All three versions are presented in Volume 3. The standard and expanded versions are designed to collect much more detailed information than has been collected in the migration modules used in previous LSMS surveys. Including either the standard or the expanded draft migration module in a future survey will yield a rich data set that should prove very useful for comprehensive research on migration.

**CHAPTER 17: SHOULD THE SURVEY MEASURE TOTAL HOUSEHOLD INCOME?** Chapter 5 (and indeed this entire book) argues that consumption is the best monetary measure of welfare in multi-topic surveys. This implies that the consumption module is essential and must always be included. In contrast, the book takes the view that collecting the data needed to calculate total household income is optional, which implies that the household enterprise, agriculture, and savings modules (which collect much of the income data) can be substantially reduced or even omitted, depending on the level of interest in these topics. Chapter 17 reviews the advantages and disadvantages of collecting the data needed to calculate total household income, and describes the circumstances under which measuring total income should be an objective of a multi-topic survey.

**CHAPTER 18: HOUSEHOLD ENTERPRISES.** Small businesses owned and operated by households are quite common in developing countries, yet it is difficult to collect accurate data on these income-generating activities. Based on their extensive experience of analyzing data from past household surveys—both LSMS and others—the authors introduce three versions of the household enterprise module (presented in Volume 3). Survey designers should choose the version that best matches policymakers' level of interest in household enterprise issues. In previous LSMS surveys, data on employment in these enterprises were collected in the employment module, but this chapter recommends that such information be collected in the household enterprise module; each of the modules introduced in this chapter collects such data. One consequence of this is that the standard version of this module is now longer than the version typically used in previous LSMS surveys, and the expanded version is even longer than the standard version.

**CHAPTER 19: AGRICULTURE.** Collecting accurate and comprehensive data on agricultural activities is difficult in any survey, and past LSMS surveys have experienced many problems in collecting such data. This chapter introduces short, standard, and expanded versions of the agricultural module (presented in Volume 3) that are very different from the agriculture modules used in previous LSMS surveys. In the standard and expanded versions, information on land owned and crops produced is gathered on a plot-by-plot basis,

rather than at the level of the whole farm as was done in previous surveys. Information on household members' work on their own farms is now obtained in the agriculture module, rather than in the employment module as in previous LSMS surveys. The short module is new, and is limited to collecting information on the households' agricultural assets and on the total amounts of each crop produced by the household.

**CHAPTER 20: SAVINGS.** It is difficult to collect data on household savings because many households are reluctant to provide savings-related information. Several previous LSMS surveys have collected a modest amount of data on savings, but these data have rarely been used in analysis due to doubts about their accuracy. Chapter 20 provides an extensive review of research on savings in developing countries, emphasizing the difficulties involved in doing such research. The two versions of the draft module introduced by this chapter (and presented in Volume 3) include several modest improvements to the module used in previous LSMS surveys. Neither of these versions is much longer than the savings modules of past surveys.

**CHAPTER 21: CREDIT.** This chapter emphasizes that to capture all of the sources and uses of credit in a way natural to respondents, questions on household credit use must be inserted in several of the survey's modules. Such questions should be inserted in the housing, consumption, household enterprise, and agriculture modules, as well as in the community questionnaire and in a special credit module. In contrast with most past LSMS surveys, the draft credit module introduced by this chapter (and presented in Volume 3) gathers information at the level of the individual rather than at the level of the household.

**CHAPTER 22: TIME USE.** LSMS surveys have traditionally included neither comprehensive measures of time use nor modules dedicated to time use. This chapter discusses the experience of special time-use surveys, and uses this experience to formulate a special time-use module (presented in Volume 3). This module will be of particular interest to researchers concerned with intrahousehold issues. However, the draft module is lengthy and could crowd other modules out of the survey (since there is a limit to the amount of time households are willing to be interviewed). Further experience will be needed in implementing such a

module as part of a multi-topic survey and in analyzing the data collected before it becomes clear whether to routinely include time use questions in LSMS and similar multi-topic surveys.

#### **Part 4**

This part, in Volume 2, presents four chapters that discuss several general survey design issues. These chapters are useful for survey designers to read alongside the chapters in Parts 2 and 3 that interest them.

**CHAPTER 23: RECOMMENDATIONS FOR COLLECTING PANEL DATA.** This chapter reviews the advantages and disadvantages of collecting panel data in developing countries, along with past experience of collecting panel data. The chapter recommends that panel data be collected in most surveys, provided that in successive rounds the original sample of households is supplemented with a sample of households living in dwellings that have been constructed since the first survey. This is necessary to ensure that the sample remains nationally representative when each survey is implemented. Chapter 23 also recommends that information be collected from households in the first survey that will help interviewers find these households in subsequent surveys, even when it is not certain that later surveys will attempt to collect panel data.

**CHAPTER 24: INTRAHOUSEHOLD ANALYSIS.** The study of the allocation of resources and responsibilities within households has grown in the economic literature over the last few years, and such issues are increasingly arising in policy discussions. This chapter explains which kinds of data should be collected at the individual level rather than at the household level in order to support intrahousehold analysis; from this perspective the chapter provides a critique of the modules proposed in Parts II and III. For modules deemed inadequate for intrahousehold analysis, the chapter proposes ways to modify them so that they better support such analysis. The chapter accepts that it is not feasible to collect individual-level data on all topics in an LSMS survey. Nevertheless, future LSMS surveys can be designed to support substantial intrahousehold analysis. Much of the data collected in past LSMS surveys—on employment, health, education, anthropometrics, migration, and fertility—have long been collected at the individual level. And the draft agriculture, household enterprise, credit, and miscellaneous

income modules presented in Parts II and III of this book recommend the collection of more individual-level data than were collected in previous LSMS surveys. In addition, the draft time-use module introduced in Chapter 22 is a new tool for gathering data that are crucial for intrahousehold analysis.

**CHAPTER 25: QUALITATIVE DATA COLLECTION TECHNIQUES.** Previous LSMS surveys have focused almost exclusively on collecting quantitative data, making very little use of qualitative data collection methods. This regrettable tendency probably reflects the quantitative backgrounds of most survey designers. Chapter 25 explains ways in which qualitative research methods can usefully and effectively complement quantitative data collection. The chapter concludes that qualitative methods should not be combined with quantitative methods into a single survey; instead, both methods should be used in separate but complementary data collection exercises. Quantitative surveys can benefit from qualitative methods in several ways. For example, qualitative research can be used to help survey designers formulate the exact wording of particular questions, and qualitative methods are useful for creating hypotheses about household behavior, which can then be tested using quantitative data.

**CHAPTER 26: BASIC ECONOMIC MODELS AND ECONOMETRIC TOOLS.** This chapter gives non-economists some basic information on economic models of household behavior, and reviews econometric methods commonly used in analyzing the policy questions discussed in this book. Chapter 26 is a useful reference for non-economists as they read other chapters. A

glossary at the end of Chapter 26 defines the economic and econometric terms used in many chapters of the book.

### Note

The authors would like to express their gratitude to Jere Behrman, Lawrence Haddad, Courtney Harold, John Hoddinott, Alberto Martini, Raylynn Oliver, Kinnon Scott, and Diane Steele for comments on previous drafts of this chapter.

### References

- Davies, David. 1795. *The Case of Labourers in Husbandry*. Bath, United Kingdom: Printed by R. Cruttwell for G. G. and J. Robinson.
- Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore, Md.: Johns Hopkins University Press.
- Deaton, Angus, and Salman Zaidi. 1999. "Guidelines for Constructing Consumption Aggregates for Welfare Analysis" World Bank, Development Research Group, Washington, D.C.
- Eden, Frederick. 1797. *The State of the Poor*. London: Davis.
- Grosh, Margaret, and Juan Muñoz. 1996. *A Manual for Planning and Implementing the Living Standard Measurement Study Survey*. Living Standards Measurement Study Working Paper 126. Washington, D.C.: World Bank.
- Ravallion, Martin, and Shaohua Chen. 1997. "What Can New Survey Data Tell Us about Changes in Distribution and Poverty?" *World Bank Economic Review* 11 (2): 357–82
- . 1998. Personal conversation updating Ravallion and Chen 1997. November 2 and 3. Washington, D.C.
- Ravallion, Martin, Guarav Datt, and Dominique van de Walle. 1991. "Quantifying Absolute Poverty in the Developing World." *Review of Income and Wealth Series* 37 (4).



# 2 Making Decisions on the Overall Design of the Survey

*Margaret Grosh and Paul Glewwe*

Comprehensive, multitopic household surveys such as LSMS surveys usually consist of three separate questionnaires: a household questionnaire, a community questionnaire, and a price questionnaire. Each questionnaire is composed of modules, sections that collect information on a specific topic. Questionnaires and their modules can be combined in a variety of different ways to create a multitopic household survey. There is no single right way to combine modules and questionnaires into a survey; each way has advantages and disadvantages. The key is to choose a design that provides the best fit given the objectives of, and constraints on, the proposed survey.

The starting point for designing the modules and questionnaires to be used in a multitopic survey is a set of policy questions. The overall objective of each survey is to collect the data needed to answer these questions. There are five steps involved in survey design. The first step is to define the fundamental objectives of the survey and to choose an overall survey design that best fits these objectives. This is usually done not by a single individual but by a team of survey designers who consult extensively with a broad range of individuals and organizations interested in the survey. In choosing the overall design, the team must take into account several important factors including the capacity for collecting data within the country, the funding available, and the amount and quality of data available from other sources.

The second step involves deciding which modules to include in the survey, specifying the objectives of each module, and proposing an approximate length for the modules. This step is needed because a survey that attempts to include all possible modules will be too large and complex to implement.

The third step is to work out, question by question, draft questionnaires for each module that will be included in the survey. This can be done by drawing on the detailed recommendations in the chapters in Parts 2 and 3 of this book, as well as the draft modules included in Volume 3. The fourth step is to compare the modules to each other to ensure that they are consistent and well integrated, and to combine them into draft household, community, and price questionnaires (in some cases omitting the community questionnaire). The fifth and final step is to translate and field test the draft questionnaires. Translation may not be necessary in some countries; field testing is always essential and must not be done quickly or superficially. The first two steps are discussed in this chapter. The third, fourth, and fifth steps are discussed in Chapter 3.

While these five steps should ideally be done in the order given above, in reality there is likely to be substantial movement backward and forward among the various steps. Some objectives originally set out for the survey may prove impossible to achieve given the constraints. And discussion of the detailed objectives

for each module may cause the survey design team to reassess the overall objectives of the survey. In other words, it may be necessary to take one or two steps backward at some points in order to continue to move forward. This is to be expected and even encouraged. As more is learned about what can and cannot be done, survey designers are more likely to produce a survey design that meets their objectives—which may also have become more realistic. It is better to pare down the number of objectives in order to achieve some of them than to attempt to do too much and, as a result, achieve few or none of the original objectives.

The first four sections of this chapter cover the first step of the survey design process. The first section provides an overview of who should be involved in designing and assembling the questionnaires. The second section discusses the main factors that survey designers should take into account when choosing among survey design options. The third section outlines “core” elements that must be included in any LSMS or similar multitopic survey and reviews several classic survey designs, each of which supplements the core in a different way. The fourth section presents guidelines for choosing the survey design most appropriate for each of a range of different circumstances.

The fifth and final section explains the second step of the survey process: choosing the modules to include in the survey, setting objectives for each module, and setting the approximate length of the modules.

### **Organizing a Survey Design Team**

The most important factor ensuring the success of a multitopic household survey is the involvement of the right people in the process. Designing the survey questionnaires is much too large a task for one person. Instead, a team of experts must be involved, including members of the organization implementing the survey as well as research analysts from other institutions. If the team does not contain a sufficient diversity of experts, this can have negative repercussions for the data (Box 2.1). The design team must work together with policymakers and program managers to define the overall objectives of the survey and to settle on many details at each step of the survey design process.

#### *Researchers and Policy Analysts*

It is essential to involve researchers and policy analysts in questionnaire design. This book was written prima-

#### **Box 2.1 The Importance of a Well-Rounded Design Team**

An effective survey design team must include researchers and policy analysts, policymakers, and staff from the organization implementing the survey. The problems that can arise when one or more of these groups is not involved in designing the questionnaires are illustrated by the experience of an LSMS survey implemented in Jamaica. The household questionnaire for the first Jamaican Survey of Living Conditions (implemented in 1988) was designed primarily by international experts who had little knowledge of Jamaican social programs. Although the household questionnaire was largely successful in accomplishing its analytical objectives, it had two serious flaws. First, although food subsidy policies were an important issue at the time, the consumption module did not clearly distinguish expenditures on key subsidized food items from expenditures on similar nonsubsidized items. This made it more difficult to study the incidence of food subsidies. Second, the questionnaire asked respondents about their receipt of food stamps during the previous month even though food stamps were provided only every two months. This made it difficult to identify which households had received food stamps, thus hindering the study of another issue important at the time. Fortunately, these flaws were identified and corrected in the following year's household questionnaire.

rily by researchers and policy analysts, and much of the success of past LSMS surveys in supporting policy-relevant research is due to the fact that the surveys used questionnaires designed by people who would be actively involved in the analysis of the data. Researchers and policy analysts can ensure that the information collected in multitopic surveys is well suited for policy research.

The lead role in designing the questionnaires of an LSMS or similar multitopic household survey should be given to a small group of researchers and policy analysts who share two characteristics. First, they should know what issues are of most concern to the country's policymakers. Second, they should have experience in using data from similar surveys to analyze these issues. The group of researchers and policy analysts should include members of the national planning agency, representatives from the national statistics agency, local academic researchers, and one or more people who have helped analyze or design multitopic surveys in other countries.

The team must include individuals with extensive experience in implementing and analyzing other

household surveys in the country in question. Ideally, local researchers and policy analysts should take primary responsibility for designing the survey, because they have an intimate knowledge of the country's culture, economy, and society, and they are very familiar with existing programs and key policy issues. Local researchers and policy analysts are also likely to know about previous surveys done in the country that have covered some of the topics included in the new survey. And they will know which people and institutions should be consulted during the survey design process.

It may also be desirable to involve international researchers in the design of the questionnaire, especially in countries where local data analysts are not familiar with LSMS and other multitopic household surveys. International researchers can contribute their experience about what has and has not worked in surveys in other countries.<sup>1</sup> Judicious use of the advice of both local and foreign experts will significantly improve survey design.

Past LSMS surveys have probably made insufficient use of the knowledge available from local researchers and policy analysts. Too often the involvement of local professionals has been limited to statisticians from the statistical agency (data producers) and thus failed to draw on the expertise of social policy researchers from the government or academia (data users). Statisticians may have only a limited knowledge of sectoral policy issues and programs. While they do have an important role to play, their input must be combined with input from data users to set priorities among the different possible objectives for policy research.

#### *Policymakers*

When defining the fundamental objectives of the survey, the team responsible for drafting questionnaires must seek extensive input from policymakers and program managers in the country being surveyed. The team's initial discussions with policymakers should focus in broad terms on the most important issues to be covered, which will determine the relative size of the different modules in the questionnaires. After this round of discussions, further discussions should be held to identify the important issues within each sector. Since drafting the module or modules for each sector requires a substantial amount of knowledge about how specific programs work, technical experts in many program agencies must be consulted. These

people should be consulted before the modules are created, and they should also be shown draft modules to elicit comments.

Unfortunately, in many previous LSMS surveys the survey design team did not give enough attention to communicating or consulting with policymakers. Policymakers, who are often unfamiliar with household surveys, may find it difficult to read complicated questionnaires or to imagine what analyses the resulting data could support. One option is to show policymakers and program managers examples of the kinds of tables and analyses that could be produced using data from the questionnaire; these might be either hypothetical tables for the country of the survey or tables made in other countries using data from similar surveys. Another strategy is to show policymakers a report based on the first year's data; this is an excellent way to obtain policymakers' feedback on the design of follow-up surveys. A third strategy is simply to ask policymakers what they need to know to implement effective policies.

#### *Data Producers*

It is critical that the survey design team include staff from the organization implementing the survey. This should ensure that the questionnaires designed are workable. Often data collection can be greatly simplified by making minor changes in the layout or flow of a questionnaire, changes that do not diminish the questionnaire's analytical content. Data producers are an excellent source of suggestions for such changes. They are usually also experienced in details of designing a questionnaire, such as questionnaire formatting. For all of the above reasons, the team members from the organization implementing the survey should help design, or comment on, every draft of the questionnaire.

It is also useful for the survey design team to solicit the input of experienced field supervisors, who will notice whether the instructions to the interviewer are clear, whether the skip codes are correct, and whether the format is consistent. There is of course a natural tension between data analysts, who want comprehensive information, and field supervisors, who are likely to see all of the disadvantages but few of the advantages of administering a lengthy, complex questionnaire. Each side must be prepared to make compromises and carefully listen to the other side's point of view.

## Factors for Deciding among Various Survey Designs

After the members of the survey design team have been selected, work can begin on designing the survey. The first task for team members is to review the factors that influence the overall design of the survey. This section discusses those factors in detail.

The appropriate design of a household survey or sequence of household surveys differs from country to country. The most important factors for determining the design of a proposed survey are: the kinds of policy issues the survey aims to address; the information available from existing surveys and other data sources; the country's institutional capacity for collecting data; and the financial and other resources available for implementing the survey, including any constraints on how these resources can be used.

### *Policy Issues*

The design of a household survey should reflect the policy issues it is intended to address. One way to classify policy issues is in terms of their subject matter, such as health, education, employment, or housing. Another way to classify policy issues is in terms of the kinds of data used to address them. The four most common kinds of household survey analysis used to address policy issues are: simple descriptive statistics on living standards; monitoring poverty and living standards over time; describing the incidence and coverage of government programs; and measuring the impacts of policies and programs on household behavior and welfare. This subsection reviews these four types of analysis and provides a practical example of how the information needed affects the design of the survey.

**SIMPLE DESCRIPTIONS OF LIVING STANDARDS.** The most straightforward objective for a household survey is to describe the living standards of the population at one point in time, often with particular emphasis on the living standards of the poor. This can be done by using the data to tabulate means and frequencies of key variables. The results of these tabulations are often disseminated by the national statistical agency in the form of statistical abstracts (reports) that contain a large number of tables and a minimal amount of descriptive text. It is also possible to produce more structured descriptive analyses that supplement household survey data with information from other sources.

Structural analysis of descriptive data can sometimes be used to draw conclusions about the likely impact of government policies on living standards. Examples of such analyses are the "poverty profiles" typically provided in the World Bank's poverty reports.

In both types of descriptive analysis, the range of variables used to measure living standards can vary widely; variables may be used from virtually all of the survey modules or from only a small subset of modules. In general, most of the variables included in statistical abstracts and descriptive analyses come straight from the questionnaire (for example, percentage of households that have electricity) or require only a small amount of manipulation (for example, nutritional status as derived from weight and height data). Only one "complex" variable needs to be constructed: total household consumption. Other complex constructed variables, such as total income or net wealth, are used less often in simple descriptive presentations.

**MONITORING POVERTY AND LIVING STANDARDS.** The descriptive analyses discussed above focus on living standards at one point in time. However, another important role of multitopic household surveys is to monitor how living standards and poverty change over time. When data are used for this purpose they must be comparable over time; for this to be the case, the data must be gathered using the same methods each time the survey is implemented. One aspect of such consistency concerns the design of the sample, which in each case must use the same definitions of basic concepts such as the distinction between urban and rural areas. A second requirement for comparability is that the questions defining variables of interest must remain the same each time the survey is administered. This is necessary because seemingly innocuous changes in the wording of questions can lead to serious comparability problems; changing the recall period for food expenditures can make it impossible to compare estimates of poverty and inequality over time.

Another issue to consider when monitoring poverty and living standards over time is the frequency with which indicators of living standards must be monitored. Indicators that are fairly stable over short periods of time—such as fertility and adult literacy—need not be measured each time the survey is done. However, indicators that can change more quickly, such as consumption expenditure, children's nutrition-

al status, and employment status, should be measured every time the survey is implemented. Surveys that monitor poverty and living standards over time are typically fielded every year, although it is also possible to field them biannually or semiannually.

**EXAMINING THE INCIDENCE AND COVERAGE OF GOVERNMENT PROGRAMS.** Data from multitopic household surveys can also be used to measure the incidence and coverage of specific government programs. For example, data on the enrollment of household members in public schools are useful for investigating which children benefit from the provision of public schooling. Household survey data can also be used to study participation in government assistance programs such as food stamps, cash assistance, and school meals. Another example is descriptive statistics on purchases of subsidized food items, which can be used to examine whether the benefits of specified food subsidies vary by households' levels of income. The incidence and coverage of these different kinds of programs are easy to calculate and useful for policymakers to know.

A moderate sample size (2,000 to 5,000 households) should be sufficient to evaluate programs that affect a large proportion of the population. Evaluating programs that serve a small proportion of the population usually requires using a much larger sample of households or including a disproportionately large number of target and beneficiary households in the sample.

**ESTIMATING THE IMPACT OF POLICIES ON HOUSEHOLD BEHAVIOR AND WELFARE.** Policymakers are often faced with questions that can be answered only by analyzing household behavior. Policymakers may want to know how changes in commodity taxes or subsidies would affect agricultural production or the consumption of basic food items. Answering such questions requires calculating price elasticities and thus modeling households' production and consumption decisions. Such modeling requires data that go well beyond measurement of living standards indicators.

In multitopic household surveys that attempt to model household behavior, each module that collects data on a behavior of interest is usually designed to gather information that can be used to estimate the impact of several different policy changes. The chapters in Parts 2 and 3 of this book discuss each module

in great detail and provide many examples of issues that require the modeling of household behavior. The following questions give an idea of the range of policy issues that can be addressed: What is the impact of charging user fees at government health clinics on the use of those clinics by adults and by children? How can the government encourage parents to enroll their children in school? What are the impacts of women's employment opportunities on their fertility? How do changes in prices brought about by structural adjustment programs affect the welfare and productivity of agricultural households?

**EXAMPLE SCENARIO: THE INCLUSION OR EXCLUSION OF THE ANTHROPOMETRY MODULE.**

The decision about including an anthropometry module demonstrates how the analytical potential of a multitopic household survey is related to its content. (See Chapter 10 for detailed information on the collection of anthropometric data.) If an anthropometry module is not included, the survey is not useful for studying nutrition issues. However, by collecting limited anthropometric data such as the height and weight of children under five years of age, the survey will allow analysts to describe the extent and patterns of malnutrition in early childhood. If the country studied has adopted large-scale food distribution or subsidy programs, the data can also be used to assess how well these programs are targeted to undernourished children. If the survey is repeated annually or biannually, it becomes possible to monitor changes in both the nutritional status of the population and the targeting of the program. Thus the data collected from a limited anthropometry module can address, at least partially, three of the four types of policy questions outlined above.

A full version of the anthropometry module would collect data on height and weight for all household members, not just children. Such data could be used not only to gauge adult health but also to analyze the impact of government policies on household welfare and behavior. Suppose policymakers want to predict the impact of food programs on children's nutritional status. This requires estimating the determinants of child weight and height. Because heredity is so important, parental height and weight information are needed to estimate these relationships accurately; lack of data on parents' weight and height could lead to estimates that suffer from omitted variable bias and thus do not accurately show the impact of the food

programs on children's nutritional status. In general, including the full version of the anthropometry module in the survey—measuring both adults and children—greatly expands the possibilities for examining the impact of government policies on household behavior and living standards.

Defining the objectives of a survey is often a less tidy process than the discussion so far has implied. Institutions, and people within institutions, may have varying objectives. Each sectoral ministry in a country is likely to be primarily interested in its own subject. The government as a whole may want the surveys to measure or monitor only a few indicators of welfare, while academics in the country's universities and other research institutions may want the surveys to yield the detailed data needed to model household behavior. If international agencies are financing the survey, they may have still another set of objectives. For example, they may wish to ensure that the data are comparable with similar data from other countries or that the data can be used to study issues of interest to the development community in general, even if these issues are not a high priority in the country of the survey. Whatever the objectives envisaged when the survey is first designed, it is likely that researchers will later use the data for other analytical purposes.

The multiple (and sometimes competing) objectives of household surveys are to be expected and even encouraged, since each of the groups with a stake in these surveys has its own legitimate priorities. The task of survey designers is to accommodate the different objectives as much as possible without compromising the quality of the survey.

#### *Other Information Available and Its Relation to Survey Objectives*

No household survey takes place in a vacuum. In most countries there are several other household surveys that have gathered or will gather information on issues that the new multitopic survey is intended to cover. The extent to which data from these sources influence the design of the new questionnaire depends on the amount and type of data available and on the objectives of the new survey.

If the main objective of the new survey is to describe various aspects of the living standards of the population, it may seem that the topics already covered in other surveys need not be included in the new multitopic survey. For example, if the only goal of the new

survey is to describe living standards and recent anthropometric data are already available from another survey, it may seem reasonable to drop anthropometric measurements from the new survey. However, there are two important advantages to collecting anthropometric data in the new survey. First, collecting these data would make it possible to produce descriptive tables that show simple relationships between nutritional status, as revealed by anthropometric measurements, and other variables of interest—for example, household expenditure levels. Second, collecting anthropometric data in the new survey would ensure that the anthropometric data used the same definitions and classification schemes as other survey data, and thus could be used to draw effective comparisons. If the two surveys classified, say, education levels or rural and urban areas differently, this would make it difficult to present analyses from the two surveys side by side in ways that would be simple to interpret. Analysis based on combining results from separate surveys will usually be more difficult, and thus more prone to error, than analysis based on data that have all been collected in a single survey.

The case for collecting anthropometric data is even stronger if the purpose of a new survey is to investigate the impact of nutritional status on other socioeconomic outcomes (such as education, fertility, or labor force productivity). This objective implies that the survey must include an anthropometry module, even if recent information on nutritional status is available from other sources. To conduct these kinds of analyses, the variables of interest must all come from the same household survey.<sup>2</sup>

Although it is essential that data on key household-level variables come from the same households, it is often useful to supplement household survey information using data from a source other than a multitopic survey. In some cases, price data collected for generating consumer price indices can replace the price questionnaire typically used in LSMS and other multitopic household surveys (see Chapter 13). Other such alternative data sources are time-series data on weather and maps of soil quality and topography, all of which can be used for analyzing agricultural issues. In health and education, further possibilities arise for matching household survey data with data from other sources; some countries collect data from health clinics and schools that may be matched with the communities covered in a household survey.

However, survey designers should exercise caution when contemplating this approach. Although matching data from different sources appears simple, it is often very difficult in practice. Many of the chapters in Parts 2 and 3 of this book discuss the potential for matching household survey data to data from other sources.

An important question that often arises when planning a multitopic survey is whether such a survey can replace one or more existing surveys and thereby reduce total costs without any loss of information. A multitopic survey with an anthropometric module could replace a periodic anthropometric survey primarily intended to measure the extent of malnutrition. However, a multitopic survey cannot replace all other household surveys. Labor Force Surveys often require much larger sample sizes and more frequent data collection than would be appropriate for a multitopic survey. And specialized surveys, such as Demographic and Health Surveys and comprehensive farm management surveys, contain much more data on those topics than can usually be collected in LSMS and similar multitopic surveys. Still other surveys, such as farm surveys and small business surveys, have very different sampling frames since they are based on samples of farms or businesses rather than samples of households.

A final issue is whether survey designers should implement an entirely new survey, modify an ongoing survey, or find creative ways to analyze existing data. Two arguments support implementing a new survey. First, past surveys may not have been adequately documented, or access to their data may be restricted. Second, inter-agency rivalry, arguments concerning ownership of survey data, and coordination problems when different surveys are carried out by different agencies may make it easier to begin a new survey instead of using existing data or modifying an existing survey. On the other hand, survey designers should at least consider trying to remedy these problems so that existing surveys can be used (perhaps with some modifications) to meet the designers' data and policy objectives—thus avoiding any unnecessary duplication. Examples are given later in this chapter of countries in which an existing survey was modified to be more like an LSMS survey.

#### *Institutional Capacity*

Decisions about what kind of survey to implement also depend in part on the institutional capacity for

collecting data in the country undertaking the survey. Because maintaining data quality becomes more difficult when surveys become more complex, this capacity should be carefully assessed when planning LSMS or similar multitopic surveys, which can be very complicated. In countries where the capacity to collect data is weak, it may be better to implement a limited multitopic survey yielding reliable data on a relatively small number of topics than an overly ambitious survey that could yield unreliable data on a wide variety of topics.

A survey containing 10 modules is easier to plan and implement than a survey containing 15 or 20 modules. The fewer the modules, the less time is needed by survey planners to contact different sectoral agencies and thus the less time is needed to build consensus. Also, smaller questionnaires require less time to design, and less time to carry out the fieldwork, enter the data, and manage the database. However, other steps in developing and implementing a household survey, such as planning the sample design, do not vary with the size of the questionnaire. Therefore, a survey with a questionnaire half the size of the questionnaire for a full multitopic household survey will involve substantially more than half the effort required for a full survey.

Despite the complexities of full multitopic surveys, some very successful multitopic surveys have been carried out in countries with very limited data collection capacity. Several steps can be taken to overcome the problems posed by limited capacity. For LSMS surveys, international experts have been brought in to draw the sample, draft the questionnaires and interviewer manuals, and write the data entry program. Such experts initially substitute for government agency staff, but they can also train agency staff to take their place in future surveys. It is highly recommended that countries with limited capacity for collecting survey data use such expert assistance.

In countries with weak institutional capacity, serious consideration should also be given to improving that capacity; capacity building yields long-term benefits that gradually reduce the need to use international experts to help with data collection. Genuine capacity building takes time, money, and political and managerial effort. An international sampling expert may be able to design and draw a sample for a survey in a few days, but it will usually take him or her much longer to teach local staff how to do so. Training to

build capacity requires significant resources beyond those already budgeted for a survey. Whether building a country's data collection capacity is important enough to warrant committing these resources will vary from country to country. Where capacity building is deemed necessary, the survey's work plan and budget must both be significantly enlarged.

If capacity building is a goal, a program of annual (or biannual) surveys will work better than a program for a single survey or for a sequence of surveys that take place every three to five years. An annual survey usually has a permanent allocation of skilled staff, staff time, and equipment. Even when the team works only part of the year on the multitopic survey, staff have a chance every year (or every two years) to use the skills that they have acquired in managing such a survey. And as the staff of the agency develop survey management skills, the need for technical assistance from international experts should diminish. When some staff members working on the survey leave, their replacements can learn their jobs from other staff members who have worked on earlier rounds of the survey. In addition, the continuity provided by an annual survey may make it easier to improve survey quality; if one year a problem arises in data collection or initial analysis, the people who deal with the problem are likely to be involved in planning the next survey and can better address the problem in the next survey.

In contrast, a survey carried out every four or five years may require new skills, staff, and equipment each time it is implemented. In the intervening period, many of the individuals who carried out the first survey may have moved on to other jobs either inside or outside the statistical agency. Those who remain may not have been involved in planning the previous survey, and the skills of those who were involved may have deteriorated over time. Vehicles and computers used in the first survey will have been allocated to other purposes, and some may have ceased to function altogether. Most importantly, much of the institutional memory about problems and potential solutions may have been lost.

A final note of caution is needed regarding institutional capacity. Sometimes, even when a statistical agency has sufficient management and technical capacity to implement a complex multitopic survey, there may not be enough experienced supervisors, interviewers, or data entry operators. Lack of data

entry operators is not a serious problem since they can be trained in a matter of weeks, and no previous experience is required. However, it takes longer to transform government staff with no household survey experience into competent interviewers. While interviewers may be trained in a month, it is not so easy to compensate for little or no interviewing experience. Experience is even more important in the case of supervisors. It may take years to overcome shortages of experienced interviewers and supervisors.

#### *Constraints Imposed by Funding Sources*

Surveys are always constrained by their funding. Most LSMS and similar multitopic household surveys receive some portion of their financing from sources other than the national budget, at least initially. As a result, they are subject to constraints associated with both the national budget and funds from external sources.<sup>3</sup>

The first and most obvious constraint imposed by the source of funding is the total amount of funds available. National budgets are often very restricted. Some external funding sources have upper limits for how much may be spent on a single project, and most have administrative procedures that grow in complexity as the size of a project increases. Also, the larger the survey budget, the more difficult it is for survey planners to justify using the money for the proposed survey rather than for some other purpose. Limitations on the size of the budget often constrain the size of the sample used in the survey and in some cases curtail the survey's analytical depth and breadth.

Another potential constraint relates to the time period over which funds may be spent. Funding agencies may stipulate that a survey project be completed in only one or two years, even though a single full-scale survey can easily take three years or more to complete—6 to 18 months to plan, a year for fieldwork, and 6 to 18 months for data dissemination and analysis. Moreover, chances of obtaining future funding can influence whether a proposed survey is carried out only once or is the first of a series of surveys. And funding limitations can affect such other aspects of the survey as the thoroughness of the survey designers' work during the planning stage, whether the fieldwork is spread over a full year or concentrated into a period of a few weeks, and the amount of analytical work funded from the survey project's budget.

Finally, many funding agencies also have rules on how survey funds can be spent. These rules may impose controls on: the percentage of funds spent in local or international currency; the balance between recurrent and investment costs; the amounts that can be spent on the salaries of local staff, survey equipment, and payment of international experts; the nationalities of such experts; and various aspects of budgeting, accounting, and procurement. Spending rules rarely influence big issues of survey design (such as survey duration, sample size, or questionnaire design), but they can affect many details of the structure and implementation of a project. Rules that prohibit the shifting of expenditures between items or between time periods may limit the ability of survey planners to deal with unanticipated problems. For example, an additional international expert may be needed quickly, but may be difficult or impossible to obtain because hiring this expert was not included in the original budget. The end result can be a delay in the survey or a reduction in quality. Another example is if an accident occurs involving a survey vehicle; fieldwork may be delayed if expenditures to repair or replace the vehicle cannot be made available promptly.

### Summary

The analytical objectives of a survey, the availability of information from other sources, local institutional capacity, and constraints imposed by funding are all key factors that typically affect whether to perform a new survey and the form such a survey will take. Many other factors are also critical, including institutional inertia and rivalry and the compromises required to build a coalition to support and conduct a survey. However, it is difficult to provide general advice because these factors usually depend on the setting of the survey; survey planners must deal with these issues as best they can given the particular circumstances they face.

### Classic Survey Designs

There are three basic ways to combine modules into questionnaires and combine those questionnaires into a survey or sequence of surveys: the full LSMS-type multitopic survey, the scaled-down LSMS-type survey, and the core and rotating module survey. All of these survey formats must include certain “core” components. This section outlines the core survey compo-

nents, discusses the strengths and weaknesses of each of the three main survey types, and describes two other survey options.

### The Core

Any LSMS-type multitopic survey must collect certain essential information about the household, its members, and the local community, including:

- A roster that lists, and collects basic information about, all household members.
- Detailed information on household consumption expenditures.
- Basic housing data such as type of dwelling, water source, type of toilet, and whether the dwelling has electricity.
- The education of all household members, including who is currently in school.
- The employment status of everyone of working age and, for those who are working, their occupation, the number of hours they worked during the previous seven days, and, if they are employees, their wage earnings.
- The receipt of money or in-kind assistance from key government or NGO programs.
- The use of social services and programs, such as government health facilities, schools, agricultural extension services, and social assistance programs.
- Basic information related to the design of the sample and the outcome of the household interview.
- Local prices of basic food and nonfood goods (unless price data are available from another source, or the country is so small and its markets so well integrated that there is very little regional price variation).

These components are referred to in this book as the *essential core*. In addition to the essential core, it is *highly recommended* that the following five types of information be collected in LSMS and similar multitopic surveys:

- Anthropometric measurements (height and weight) of children 0–5 years old (unless malnutrition is known to be negligible in the country).
- The immunization status of children 0–5 years old.
- Information on basic household assets such as durable goods, housing, land, and the capital equipment used for agricultural activities and nonagricultural household enterprises.
- Information on interhousehold transfers.
- Information on rental payments for those households that rent their dwellings.

In this book the set of modules formed by adding these five components to the essential core is referred to as the *recommended core*.

The essential core of an LSMS or similar multitopic survey collects the information needed to describe poverty and to monitor it over time. The recommended core adds some very basic child health information, along with information on assets, inter-household transfers, and rental payments (the use of which will be explained below). Judgments about which data are part of the essential and recommended cores are based on many years of experience that World Bank staff have in using data from LSMS surveys to produce poverty profiles for a wide range of developing countries. Table 2.1 lists the components of both the essential and recommended cores of LSMS-type multitopic surveys. The paragraphs that follow describe each of these components in greater detail.

**HOUSEHOLD ROSTER (ESSENTIAL).** Virtually every household survey should begin by determining how many people belong to each household and collecting very basic information on each household member, including age (or date of birth), sex, nationality, relationship to the head of household, and marital status. Part A of the household roster introduced in Chapter

6 (and provided in Volume 3) collects such basic information, along with another piece of information that is less essential: questions that link each married individual to his or her spouse.

**CONSUMPTION EXPENDITURES (ESSENTIAL).** The experience of LSMS surveys and other household surveys strongly suggests that household consumption expenditures are the single most important indicator of household welfare that can be obtained from a household survey. (See Chapters 5 and 17 for further discussion on this point.) Chapter 5 describes how to collect data on consumption expenditures, stressing that there are no costless shortcuts for collecting such data. In some circumstances it might be possible to omit questions on the ownership of durable goods and on transfers given to other households, but the rest of the consumption module is an essential part of the core and should not be reduced further. Data on household expenditures on education, health, and housing are collected in the core elements of those modules (discussed below) and need not be included in the consumption module. Consumption in the form of in-kind payments (such as meals, clothing or transportation) from employers is best collected in the employment module.

**Table 2.1 The Essential and Recommended Cores of LSMS-Type Multitopic Surveys**

Module	Sections used
<b>The Essential Core</b>	
Household Roster	All of Part A except questions 8 and 9
Consumption	All questions except transfers given to other households (Part D) and ownership of durable goods (Part E)
Housing	Questions A1–A7, B1–B5, C1–C3, and C13–24 of the short module
Education	All questions in the short module
Employment	Questions A2–A13, B1, B2, B7–B10, D3, D4 and D8–D17
Transfers and Other Nonlabor Income	All of Part B1; see text for further discussion
Health	Questions 10–38 of the short module
Metadata	Household Identification and Control submodule; Questions 1–4 in Summary of Visits and Interviews submodule
Prices	30–40 food items and 10–20 nonfood items
Credit	Questions 9–14 and 21–28 of the short module (on credit obtained from NGOs or government agencies)
Agriculture	All of Part F (use of agricultural extension services), which is the same for all modules
<b>Additional components for the recommended core</b>	
Anthropometry	Entire module, for children 0–5 years old
Health	All of Part C (immunization)
Consumption	All of Parts D and E
Housing	Questions C7–C12 of the short module
Household Enterprises	Part G of the short module, questions 1–3
Agriculture	Parts A, B, and E of the short module.
Transfers and Other Nonlabor Income	Questions on income from interhousehold transfers

Source: Authors' recommendations.

**HOUSING (ESSENTIAL).** Information on housing, including the type of dwelling, the construction materials used, the number of rooms, the availability of electricity, the source of drinking water, and the type of toilet, are very basic indicators of living standards. They provide analysts and policymakers with information on a household's standard of living that goes beyond consumption expenditures. Because housing information is simple to collect, it should be included in any LSMS-type survey. The long version of the housing module introduced by Chapter 12 (and presented in Volume 3) collects substantially more housing data than are necessary for the essential core, and even the short version is longer than the essential core. Only the following questions from the short version of the housing module need to be included in the essential core: A1–A7, B1–B5, and B11–B21. Even some of these questions can be omitted in some countries. The questions on heating (B18–B21) can be removed for countries with warm climates, and the questions that distinguish between wet and dry seasons (B1–B3) can be simplified in countries where this distinction is not important.

Another part of the core data set, housing-related consumption expenditures (such as expenditures on electricity, water, and cooking fuel), are most conveniently collected in the housing module rather than in the consumption module. A final useful indicator of living standards is information on the ownership of the household's dwelling. Questions C1–C3 in the short version of the housing module collect expenditure information, and questions C13–C24 collect ownership information.

**EDUCATION (ESSENTIAL).** Education is both a determinant and a key indicator of living standards. The short version of the education module introduced in Chapter 7 (and presented in Volume 3) comprises all of the essential core questions on education. The only questions that might be omitted are the two questions on grade repetition.

The short education module assesses education from several different angles, including school attainment, current enrollment, and education expenditures. Information on the school attainment of household members ages 5 and older is easy to collect and has many analytical uses (such as classifying households in terms of the education level of their head of the household). School enrollment among children, another key

indicator of living standards, is also easy to collect. And it is usually more convenient to collect information on household education expenditures in the education module than in the consumption module.

**EMPLOYMENT (ESSENTIAL).** Basic employment information on household members of working age (7 and older in many countries) should be collected as part of the essential core of any LSMS-type survey. The most important source of income for poor people in developing countries is their labor; employment data, including information on unemployment, indicate how this labor is being used.

Essential employment information includes each person's occupation and the number of hours that he or she has worked in the previous seven days. While it would also be useful to gather data on the incomes of all employed household members, this is not easily done for the self-employed (see Chapter 17 for further discussion). However, income data should still be collected for employees even when such data cannot be collected for the self-employed, for two reasons. First, these partial data are useful for understanding which occupations pay well and which do not. Second, since data are already needed from employees on in-kind benefits provided by their employers (in order to calculate consumption expenditures), it would seem strange to ask about those benefits without first asking about money income.

The short version of the employment module introduced by Chapter 9 (and presented in Volume 3) collects more information than the core of an LSMS-type survey requires. The following questions from the draft employment module constitute the essential core: A2–A13, B1, B2, B7–B10, D3, D4, and D8–D17. Job-specific information—questions in Parts B or D—should be collected both for the person's main occupation and for any secondary occupation. (The main occupation is the job the respondent spent the most hours doing during the previous seven days.)

**GOVERNMENT AND NGO TRANSFERS (ESSENTIAL).** Many developing countries have programs that provide money or in-kind assistance to households. Some of these are government programs and some are run by nongovernmental organizations (NGOs). Examples of these programs include cash welfare payments, pensions, unemployment insurance, food stamps, food rations, school feeding programs, community soup

kitchens, scholarships, and free or subsidized textbooks. While the range of programs is very wide, there are usually only a few sizable programs in any particular country.

A key policy question that LSMS surveys can address is who benefits from these programs. However, only programs that reach a substantial fraction of the population can be studied with the relatively small sample sizes recommended for LSMS and similar multitopic surveys.

Questions about government and NGO transfer programs should not necessarily all be in the same module (a fact that makes this part of the core difficult to standardize). While questions about cash income fit best in the transfers and other nonlabor income module, questions on school feeding programs should probably be in the education module. However, Part B1 in the transfers and other nonlabor income module is a good place to start collecting this information.

**SOCIAL SERVICES (ESSENTIAL).** Related to programs that provide cash or in-kind assistance are programs that provide services. The most common examples of social services are public schools, public health services, agricultural extension services, credit programs, public work schemes, electricity supply, public water supply, and sewage systems. LSMS surveys and similar multitopic surveys should always collect some information on the use of social services, at least enough to measure variation in access to and utilization of such services across different socioeconomic groups.

As with direct assistance to households, the types of programs and the amount of detail needed to identify who benefits from them will vary among countries. School enrollment information is already collected in the core, as discussed above, although additional information may need to be collected on any school services that are available to some students and not to others, such as tuition waivers or afterschool programs for disadvantaged students. Information on the use of public health services is also very important; such information is collected in questions 10–38 of the short version of the health module (introduced in Chapter 8). Data on the use of agricultural extension services are collected in Part F of all versions of the agriculture module introduced in Chapter 19. Some countries have subsidized credit programs to assist poor households; information on these programs is collected in questions 9–14 and 21–28 of the short

version of the credit module introduced in Chapter 21. It is possible to identify beneficiaries of public works programs by adding one or two questions to the employment module that ask whether an individual's current employment is related to such a program. Finally, information on housing-related physical infrastructure services—such as water, sanitation, and electricity—is collected in the core of the housing module, as discussed above.

**METADATA (ESSENTIAL).** The last type of information that must be collected in the household questionnaire of any LSMS-type survey consists of basic data on where the household fits in the sample and on the outcome of the interview. This type of information, known as “metadata,” is discussed in Chapter 4. For the essential core, it is not necessary to collect all of the information covered in the metadata module. The essential metadata are the date of the interview or interviews, the identification (ID) codes for the household and its primary sampling unit,<sup>4</sup> the ID codes of the interviewer and the other team members who collected, checked, or entered the data from that household, information on whether an interview actually took place (and if not, why it did not), and perhaps some data on the ethnic group and religion of the household. This information is collected in the metadata module, on the Household Identification and Control submodule and in questions 1–4 of the Summary of Visits and Interviews submodule.

**PRICES (ESSENTIAL).** Price information should be collected at the level of the community (the primary sampling unit) since all households in a given community face the same prices. How to collect price information is discussed in Chapter 13. The main task is to select the items for which price data will be collected. While the exact items will vary across countries, prices should be collected for at least 30–40 of the most commonly consumed food items and 10–20 of the most commonly purchased nonfood items. In a few countries other sources of reliable price data may already exist for both urban and rural areas; if these data can be matched to the communities covered in the survey, there is no need to collect new price data. And in some small countries such as Jamaica, prices vary little among regions. In these cases, no price data need to be collected as long as national price data exist that show changes in prices over time.

**ANTHROPOMETRIC MEASUREMENTS (RECOMMENDED).**

Anthropometric data, particularly on height and weight, should be collected for children 0–5 years old in almost every LSMS or similar household survey. Stunting (low height for age) and wasting (low weight for height) are common measures of children’s nutritional status; height and weight data are critical in countries where children are at risk of malnutrition. And collecting basic anthropometric data about children is simpler and more reliable than collecting other data on health status. The details of how to collect children’s anthropometric data are explained in Chapter 10.

Collecting height and weight data requires some effort. The data are collected using special equipment that is bulky and troublesome to carry around to each household. One consequence of this is that another individual is often added to each survey field team. If collecting children’s height and weight data were easier, such anthropometric measurements would have been classified as part of the essential core of any LSMS-type survey.

**IMMUNIZATION (RECOMMENDED).** Almost all LSMS and similar multitopic surveys should collect immunization records for children ages 0–5. In recent years child immunization programs have dramatically reduced the incidence of several life-threatening childhood diseases in many developing countries—significantly reducing infant and child mortality rates. However, many countries still do not have 100 percent immunization coverage. Therefore, information on the extent of coverage and on where coverage is low is important for almost any analysis of living standards. In addition, since child immunization coverage can change dramatically over a year or two, it serves as a useful indicator of changes in the provision of government services during periods of economic or social instability. Child immunization information is collected in Part C of the health module introduced by Chapter 8.

**ASSETS (RECOMMENDED).** Household assets include information on any consumer durable goods owned by the household, the value of owner-occupied housing, and the ownership of land and capital assets related to agricultural activities and household enterprises. There are several reasons for collecting these data. First, the possession of household durable goods such

as radios, televisions, bicycles, motorcycles, and cars is a simple indicator of living standards. Second, the sum of the value of all these different household assets gives a rough (and admittedly incomplete) indicator of household wealth. Third, data on the ownership of land and on capital assets used in agricultural and nonagricultural enterprises indicate productive assets. Fourth, in some countries, particularly countries of the former Soviet Union, there is evidence that adding the consumption derived from durable goods and housing to total consumption can lead to substantial changes in the relative economic positions of different types of households.

Information on the ownership of consumer durable goods can be collected using Part E of the consumption module. Data on the value of owner-occupied housing are collected in the short version of the housing module, using (at minimum) questions C1 and C11, with C3 and C12 providing alternative valuations. A short set of questions on the assets used in household enterprises is provided in Part G of the short version of the household enterprise module; only questions 1–3 are needed. Parts A, B, and E of the short agriculture module collect a modest amount of information on households’ land holdings, livestock, machinery, and other agricultural assets.

**PRIVATE INTERHOUSEHOLD TRANSFERS (RECOMMENDED).** Private interhousehold transfers, which are pervasive in many countries, are used by many households to cope with poverty and economic vulnerability. Transfers received are covered by the transfers and other nonlabor income module (introduced in Chapter 11) and transfers sent are covered by the consumption module (introduced in Chapter 5). At least the short versions of the private interhousehold transfer submodules should be used in virtually all surveys. Even in a relatively simple survey it may be worthwhile to use the standard version of the submodule on transfers received.

**RENTAL PAYMENTS FOR HOUSING (RECOMMENDED).** Estimates of the annual rental value of dwellings are needed to estimate the consumption value of housing for households that own these dwellings. In most countries such estimates can be calculated by estimating the relationship between basic housing characteristics, which are already part of the core, and the rental payments made by households that rent their

dwellings. The key piece of information needed is the rental payments of households that rent. Questions C7–C12 in the short version of the housing module collect information on rental costs.

#### *Full LSMS-type Survey*

In practice, the essential core—and even the recommended core—will tap only a small part of the potential policy uses of an LSMS-type survey. In most LSMS and other multitopic surveys, much more can and should be added to the questionnaires to gather information beyond what is collected in the core. This subsection and the two that follow it discuss different ways to add to the core by expanding modules and combining them to form a survey or sequence of surveys.

A full LSMS-type multitopic household survey can be formed by combining the short or standard versions of most of the modules in the household questionnaire with the corresponding parts of the community and price questionnaires. This produces a household survey similar in design to the original LSMS surveys first used in 1985, except that the modules presented in Volume 3 of this book (and described in Parts 2 and 3) include revisions based on 15 years of experience with LSMS and other household surveys.

Because some of the standard versions of modules presented in Parts 2 and 3 are significantly larger than versions used in the original LSMS surveys, a household questionnaire including all of the standard modules would almost certainly be too large to be practical. Thus the household questionnaire of a full LSMS-type survey needs to be trimmed, either by replacing the standard versions of some modules with their short versions or by dropping some nonessential modules.

A well designed full LSMS-type multitopic survey collects information that measures or otherwise describes:

- Household consumption.
- Household income.
- Key nonmonetary indicators of welfare such as nutritional and health status, education status, and housing conditions.
- Many aspects of household behavior, such as income-generating activities, human capital investments, fertility, and migration.
- The local economic environment (including prices and the availability of services).

- Participation in specific government programs such as food stamps programs, job training programs, and agricultural extension services.

Having all this information for a group of households makes it possible to describe many indicators of living standards, estimate the determinants of different dimensions of living standards and different types of household behavior, and estimate the relationships between dimensions of living standards and household behavior (such as the impact of children's nutritional status on their school performance).

The full LSMS household questionnaire is long and complex. In almost all cases it is too long to be completed in a single visit by an interviewer to a household. Instead, an interviewer typically visits each household twice. All of the individual-specific modules (roster, education, health, employment, and migration) are administered in the first visit, sometimes with the addition of one or two household-level modules such as housing. The interviewer makes an appointment for a second visit, usually about two weeks later, to reinterview household members who are most knowledgeable about the other household-level modules (such as consumption, agriculture, and household enterprises). To ensure that high quality data are collected and to keep the budget within reasonable limits, the samples in full LSMS-type surveys are usually relatively small—between 2,000 and 5,000 households. Samples of this size are still large enough to provide accurate information on the nation as a whole, on rural and urban areas, and on a small number of geographic regions. However, such samples are not large enough to provide accurate statistics for each state, province, department, or district in a country. Even at the national level, they cannot provide precise information on phenomena that do not pertain to most households or individuals—such as post-secondary education or participation in a program used by only a small fraction of the population. See Grosh and Munoz (1996) for a more thorough discussion of sample size and sampling issues.

In most cases it is not worthwhile to implement a full LSMS-type multitopic survey every year. Much of the analysis for which LSMS surveys are designed does not need to be repeated annually. For example, while it is important to understand the determinants of fertility, it is unlikely that these determinants change greatly from one year to the next. Sizable changes are likely to occur only over the course of several years, as

economic conditions and people's attitudes change. Another reason not to implement a full survey every year is that it is costly to administer such a comprehensive household questionnaire, and requires substantial work at each stage. Therefore, a full LSMS-type survey should be implemented only once every three to five years.

During 1985–99 the following countries implemented full-size LSMS surveys: Algeria, Brazil, Côte d'Ivoire, Ecuador, Ghana, the Kyrgyz Republic, Mauritania, Morocco, Nepal, Pakistan, Panama, Peru (in 1985–86, 1991, and 1994), Turkmenistan, and Vietnam.

### *Scaled-down LSMS-type Survey*

A scaled-down LSMS-type survey can be constructed by omitting some modules from the household questionnaire of a full LSMS-type survey and by abridging other modules. Such a survey will still be a multitopic survey, but will cover fewer topics than a full-size survey would. Substantial reductions in the size of the household questionnaire may mean that the questionnaire can be completed in a single visit by the interviewer to the household, as compared to the two visits needed for a full LSMS-type multitopic survey.

The extent to which various modules should be reduced or eliminated will depend on which policy questions are most important in the country in question. However, there is a limit to how much the questionnaire can be cut. The essential core of an LSMS or similar multitopic survey, as described above, must remain. In addition, the elements that are added to form the recommended core (data on anthropometrics, child immunization coverage, basic household assets, interhousehold transfers, and rental payments of households that rent their dwellings) should almost always be included. The community questionnaire may or may not be included in a scaled-down survey, but the price questionnaire should always be used, except in those rare cases in which fully adequate price data already exist or price variation across regions is negligible. Overall, the analytical objectives of a scaled-down LSMS-type survey are more modest than the objectives of a full-size survey.

One common way to abridge the questionnaire is to decide not to collect the data needed to measure total household income. Not measuring total household income allows survey designers to delete most of the agriculture and household enterprise modules,

retaining only questions on agricultural extension services that are part of the essential core and questions on assets that are part of the recommended core. Yet the questions on wages from the employment module and the questions on public and private transfers from the transfers and other nonlabor income module should be retained, as they are part of the essential core of any LSMS-type survey.

Other questions that can be dropped are questions on any aspects of household behavior that are of little interest to policymakers. The savings, credit, fertility, and migration modules have often been deleted from previous LSMS surveys. Because the new time-use module is quite lengthy, it is also a candidate for omission, unless data on time use are of particular interest to policymakers. If analysts aim to measure use of social services but not to estimate the determinants of demand for them, survey designers could choose to use the short, rather than the standard, versions of the health and education modules.

An alternative way to obtain a scaled down multitopic survey is to “scale up” an existing single-topic household survey, such as a labor force or household expenditure survey. In Romania, Latvia, and Bangladesh, new modules on the use of social services and programs were added to existing household income and expenditure surveys. In Guyana, households that had been interviewed in a previous income and expenditure survey were revisited to collect information on health, education, and anthropometrics; the separate data files were later merged for purposes of analysis. In Jamaica, households from the Labor Force Survey were revisited by interviewers who administered the Survey of Living Conditions; the two data files were later merged. In Paraguay, additional modules were added directly to the Labor Force Survey questionnaire.

Scaling down the household questionnaire of a full LSMS-type survey reduces the analytical potential of data collected, especially in parts of the questionnaire that are dropped or abridged. A reduced questionnaire produces fewer descriptive statistics on many dimensions of household welfare than would be possible using a full-size survey. Data from a scaled-down questionnaire can be used to analyze only a few of the determinants of living standards. And such data reductions substantially reduce the range of analytical methods that can be used.

A scaled-down LSMS-type survey can be implemented fairly often, perhaps annually or every other

year. Such frequent implementation is desirable because one of the main uses of data from a scaled-down survey is to monitor changes in poverty and other dimensions of welfare over time. Also, the fact that a scaled-down survey collects less data on the determinants of household welfare and behavior than does a full-size survey means that implementing it frequently wastes less resources than would implementing a full LSMS-type survey every one or two years. Another advantage of a scaled-down survey is that it is easier and less expensive to carry out than a full-size survey. Finally, a scaled-down survey can be carried out using somewhat larger samples than a full LSMS-type survey because it is subject to fewer managerial and budget constraints.

Scaled-down LSMS surveys have been carried out, with World Bank support, in Albania, Azerbaijan, Bolivia, Bulgaria, Pakistan (1995/96 and 1996/97), Peru (1990), and Tanzania.

#### *Core and Rotating Module Design*

The “core and rotating module” design for a multi-topic household survey is an attempt to combine the advantages of full and scaled-down LSMS-type surveys. In this design, a scaled-down LSMS-type survey forms the “core,” while one or two modules are added or greatly expanded each time the survey is carried out. Modules that are added or expanded in any given year revert back to their “core” size the following year, creating a module “rotation” scheme for the modules that go beyond the core. In most cases the survey is fielded annually, although it can also be a semiannual or biannual survey. The core that is repeated each time the survey is implemented must include the essential core described above, and in almost all cases it should include everything in the recommended core. In many cases the core of a core and rotating module design should collect additional information as well, in order to provide a more detailed picture of household welfare each time the survey is implemented.

An example of how to implement this approach would be to use only the core in the first year of the survey, in order to focus on making sure that the core works well. In the second year the health module in the household questionnaire would be expanded to gather more detailed data on individuals’ health status and behavior, the kinds of health care sought, and the cost and quality of that health care. In addition, a health facility questionnaire could be added (see Chapter 8 for

further details on health facility questionnaires). In the third year the health module would return to its original “core” size and a new subject, such as education or savings, would be given special emphasis. Expansion of any particular module might require making some additions to other modules in the survey to ensure that the analytical potential of the data collected in the expanded module could be fully exploited. Each chapter in Parts 2 and 3 of this book explains what data are needed from other modules to complement the data collected in the module covered by that chapter.

The core and rotating module design is a hybrid of a full LSMS-type multitopic survey and a frequently implemented scaled-down LSMS-type survey. Implementing a core and rotating module survey annually would allow for the same monitoring of poverty and welfare that is possible with data from an annual scaled-down survey. In addition, in each rotation of a particular module, this kind of survey would collect the data analysts need to study the determinants of household behavior for a specific topic—in other words, data comparable to what are collected in a full-size survey. It might even be possible to use data from the scaled-down modules to study topics that are not emphasized by the survey in a particular year.

The cost and sampling implications of the core and rotating module design lie somewhere in between those of a full-size LSMS-type survey and those of a scaled-down survey. Perhaps of greatest concern in the core and rotating module design are the institutional arrangements for developing, implementing, and analyzing the special modules. While for both full-size and scaled-down LSMS-type surveys it is possible to put a lot of effort into the design of the first survey and give less attention to improving its design in subsequent years, implementing the core and rotating module design means that the questionnaire needs to be significantly modified each year—requiring much more attention from survey designers after the first year.

Indonesia’s SUSENAS is a long-standing example of a core and rotating module survey design. Jamaica’s Survey of Living Conditions, which began in 1988, was the first LSMS survey to adopt this approach. A new LSMS survey in Cambodia is just starting to develop such a system, as is the Bangladesh Household Expenditure Survey. (The Bangladesh Household Expenditure Survey is not usually regarded as an LSMS survey; however, it has adopted much of the LSMS methodology.)

### Special Purpose Sample Designs

There are two other possible survey designs, both of which use special purpose samples (that is, samples that are not nationally representative). The first is a survey that samples a special population that is of particular interest for analytical or policy purposes.<sup>5</sup> An example of this is a sample of households within a single city that is used to study issues pertaining to that city, such as the housing market, the water supply system, or urban air pollution. Two LSMS surveys of this type have been performed: one in the Kagera region of Tanzania, focusing on areas with high prevalence of AIDS, and one in rural areas of Northeast China, focusing on the agricultural activities of rural households.

A second kind of special purpose survey is one in which the sample is drawn solely for purposes of program evaluation. In this type of survey, a group is observed both before and after the benefits of a particular service or program are made available to this group. Alternatively, the sample may be composed of two groups, one consisting of the households who benefit from the service to be evaluated—the treatment group—and the other consisting of households that are similar to the first in every respect except that they do not benefit from the service—the control group.

These special-purpose samples usually gather detailed data on the topic being studied, whether it is a specific sectoral issue (such as agriculture) or a program to be evaluated. There are so many ways to design such surveys that this book cannot hope to cover all of them. However, since special purpose surveys typically collect data on many general characteristics of the sampled households (such as size, composition, living standards, labor force status, and education), designers of this kind of survey can use the modules proposed in this book as a guide for collecting this supplemental information. The experience of past LSMS surveys has been used in designing special purpose surveys to evaluate the impact of educational reforms in El Salvador. And the Nicaragua LSMS survey included a special sub-sample designed to evaluate the impact of that country's Social Investment Fund.

### Matching Circumstances and Designs

This section provides some approximate rules of thumb for choosing among the three common survey design options discussed in the previous section. These recommendations should not be thought of as rigid or

beyond debate; instead, they should be thought of as a starting point for making survey design decisions. This is the case for several reasons. First, the dividing lines between the three basic survey designs are flexible, as it is possible to develop “hybrid” surveys that merge characteristics from the different survey design options. Second, individual countries may not fit neatly into the categories implicit in the rules. Third, surveys may have multiple analytical objectives. Finally, funding constraints are not explicitly considered here.

Survey planners should consider the following general “rules of thumb” when deciding what kind of survey to implement:

1. Countries with sufficient institutional capacity to implement a complex survey should use either a full LSMS-type multitopic survey every three to five years or a core and rotating module design; both options can serve a broader range of analytical objectives than can a sequence of scaled-down LSMS-type surveys.
2. If annual (or biannual) monitoring of living standards or poverty is the most important analytical objective, either a sequence of frequent scaled-down LSMS-type surveys or a core and rotating module survey should be adopted. In contrast, a full-size multitopic survey is inappropriate because cost and efficiency considerations imply that such surveys should be implemented only every three to five years.
3. No new survey is needed if the main objective is to provide periodic descriptive information (say, every three to five years) or to examine the coverage of government programs in countries where ample data are already available from other sources.
4. If the main objective is to gather periodic descriptive information or to examine the coverage of government programs, a core and rotating module design should not be chosen. Such a design would collect data much more frequently than is necessary.
5. If the main objective is to model household behavior, either a full LSMS-type survey or a core and rotating module survey should be chosen. A series of scaled-down surveys would be insufficient for modeling household behavior.
6. If the main objective is to model household behavior and very little other data are available, a full-size multitopic survey is preferable to a core and rotating module survey since the latter cannot

supply detailed information on all topics until it has been in operation for several years. The core and rotating module design can be adopted after one or two full LSMS-type surveys have been carried out.

7. If the main objective is to model household behavior and a large amount of other data are available, the core and rotating module survey is preferable to a series of periodic full LSMS-type surveys because the core and rotating design allows poverty to be monitored more frequently over time.
8. If the institutional capacity in the country is limited and the survey aims either to monitor poverty and living standards annually or to provide descriptive information (including coverage of government programs) periodically, a scaled-down LSMS-type survey should be chosen. This survey may be either frequent (for annual monitoring) or periodic (for descriptive information every three to five years). The other options, full multitopic and core and rotating module, are too complex for countries with limited institutional capacity.

Table 2.2 summarizes the implications of these rules, showing which rules lead to which choices. Because countries with little institutional capacity cannot implement a full LSMS-type multitopic survey or a core and rotating module design on their own, they will not be able to collect data that are useful for analyzing household behavior unless their institutional capacity is either permanently improved or supplemented in the short run by using international experts. In addition, significant purchases of new equipment may be required in some countries.

### Choosing the Modules, Defining Their Objectives, and Setting Their Size

Once the basic blueprint of the survey has been selected, survey designers must decide which modules to include in the household and community questionnaires.<sup>6</sup> Designers must also define specific objectives for each module and decide on each module's approximate length. The procedures for these steps are discussed in this section. Because decisions about length and objectives ultimately depend on many country-specific details, specific recommendations cannot be provided for each possible scenario. Instead, some general guidelines and procedures are provided that should prove useful for completing this step efficiently and effectively.

Two general points must be made at the outset. First, the tasks of choosing modules, defining their objectives, and setting their approximate size are all closely related and thus must be done simultaneously rather than sequentially. The type of objectives and the number of objectives have considerable implications for the size of each module; more objectives, and more complex objectives, necessitate a larger module. Second, the objectives of each module should be consistent with the overall objectives of the survey, in terms of both the analytical objectives (describing living standards, monitoring poverty and living standards, examining the coverage of government programs, estimating the impact of policies) and the specific topics in which policymakers are interested. The overall objectives of the survey already provide some information on what the objectives of many of the modules will be.

**Table 2.2 Recommended Survey Designs for Different Settings**

Availability of other data		Analytical objective	
		Describing living standards or examining program coverage	Monitoring living standards or poverty
<b>Countries with sufficient institutional capacity</b>			
Limited	Full LSMS-type survey (Rule 1 + Rule 4)	Core and rotating module (Rule 1 + Rule 2)	Full LSMS-type survey (Rule 5 + Rule 6)
Ample	No new survey needed (Rule 3)	Core and rotating module (Rule 1 + Rule 2)	Core and rotating module (Rule 5 + Rule 7)
<b>Countries with limited institutional capacity</b>			
Limited	Periodic scaled-down LSMS-type survey (Rule 8)	Frequent scaled-down LSMS-type survey (Rule 8)	Full LSMS-type survey (Rule 5 + Rule 6) <sup>a</sup>
Ample	No new survey needed (Rule 3)	Frequent scaled-down LSMS-type survey (Rule 8)	Core and rotating module (Rule 5 + Rule 7) <sup>a</sup>

a. International experts must be hired to carry out key tasks.

Source: Authors' recommendations.

### *Choosing modules*

A good first step in choosing modules is to set the upper and lower limits of what can be included in a multitopic survey. The lower limit is the essential core discussed above; in almost all cases this lower limit should be expanded to include the additional elements that are in the recommended core. The upper limit will depend on country-specific circumstances such as the capacity of the statistical agency and the willingness of households to participate in lengthy interviews. It is never possible to include all of the modules in any one survey.

An important question to address relatively early when making decisions about modules is whether the survey will attempt to collect enough data to calculate total income. The advantages and disadvantages of collecting these data are discussed at length in Chapter 17. Clearly, if survey designers decide to collect the data needed to calculate total income, the agriculture and household enterprise modules need to be included in the questionnaire.<sup>7</sup> If designers decide not to collect income data and there is little interest in these two modules, they can be dropped, except for the questions on use of agricultural extension services that are part of the essential core and the asset questions that are part of the recommended core.

It will probably not be possible to collect total income data in a scaled-down survey because it is not feasible in a single visit to a household to collect the recommended core data plus the data from the agriculture and household enterprise modules and still have room to examine other topics. This implies that it is also difficult to collect total household income in a core and rotating module survey, except when the module featured is either the agriculture or the household enterprise module; even when one of these two modules is featured in such a survey, collecting total income data may not be feasible in some countries.

Two other specific decisions to make early in this step of the survey design process are whether to collect time-use data and whether to implement a large number of the detailed environment modules (see Chapters 22 and 14, respectively). The time-use module is very long, and as such should be thought of as an expanded module. If survey designers choose to include this module, they may have to omit several other short or standard modules. While it would be feasible to include the time-use module in a full-size LSMS-type survey, this module is probably too large

to be included in a scaled-down LSMS-type survey and would work in a core and rotating module survey only if it were chosen as the topic emphasized in a particular year. The same circumstances apply to the environmental module. The full set of these environmental submodules is equivalent to a very large expanded module, and for this reason it is difficult to imagine the full set used in a single survey. The contingent valuation modules should be used only when specific improvements in services (such as urban water supply, urban sanitation, urban air quality, or rural water supply) are being contemplated.

Even a subset of the expanded environmental modules is likely to be equivalent to a large expanded module, especially if the water, sanitation, and fuel use modules are included. This being the case, it is feasible to include a large subset of the environmental modules in a full LSMS-type multitopic survey, but only a relatively small subset can be included in a scaled-down survey. In a core and rotating module survey a large subset of the environmental submodules can be used only if environmental topics are emphasized in that particular year; for all other years only a small subset would be feasible.

At this point, it is useful to give some general rules about how much room there is for modules in different kinds of multitopic surveys. For a full LSMS-type survey, the household questionnaire should be roughly large enough to include a mixture of about 15 standard or short modules. The number of modules that can be included in a scaled-down survey is probably closer to 8 or 10, most of which have to be short versions. A core and rotating module survey lies somewhere in between but is probably closer to the scaled-down survey if only one visit is made to the household. The modules chosen must in all cases include the components of the essential core; in almost all cases the modules should also include the additional components found in the recommended core.

Using these starting points for what is feasible, the next task is to consult with policymakers at the highest level to get a detailed idea of which topics are of greatest interest to them (if this has not already been done). Policymakers need to specify which topics are of overriding concern, which are of moderate interest, which are of minor interest, and which are of little or no interest. Expanded modules, if they exist, should be used for topics of overriding interest.<sup>8</sup> Standard modules should be used for items of moderate interest.

Short modules may be appropriate for items of minor interest. Items of little or no interest need not be covered in the survey unless they are part of the essential or recommended core.

The core and rotating module survey design is inherently more flexible than other classic designs; if the core and rotating survey is implemented annually, it can cover four or five topics in great detail over the same number of years by including the expanded version of one of these modules each year. Of course, survey designers still have to set priorities about which expanded module is included in the first year, which is included in the second year, and so on.

The above paragraphs provide survey designers with a scheme for generating a draft list of the modules to be included in the survey, their approximate length, and, to an extent, the objectives of each module. Needless to say, this draft list needs to be refined. This can be done by adding two new “ingredients” to the process: discussions with policymakers who specialize in particular topics or programs, and a careful reading of the chapters in Parts 2 and 3 of this book. The task is to reconcile the specific policy questions raised by these more specialized policymakers with the feasibility of collecting data to analyze them (as discussed in detail in Parts 2 and 3 of this book) given the approximate sizes of each module as specified by high-level policymakers. This process is not simple and consequently involves a certain amount of iteration.

Unfortunately, policy issues raised by specialist policymakers often require more questions than can fit into a module of the size specified in the first draft of the modules. The choice at this point is between not including many of these policy issues in the survey and expanding the module containing these questions at the expense of other modules. A third alternative is expanding the relevant module without reducing the size of any other module, but the feasibility of this option is open to question and will not become clear until a draft questionnaire is field tested.

Given this situation and the uncertainty regarding what is feasible and what is not, survey designers should use the following procedure to reconcile the specific objectives of each module with any constraints on module size. First, designers should ask policymakers who specialize in a given topic to rank the policy issues in order of importance, so that the module can collect the data needed to analyze the most important policy issues despite the inevitable constraints.

Second, for each module, survey designers should match the policy issues raised by policymakers with the data required to analyze them, as laid out in each chapter of Parts 2 and 3. One way to do this is to choose the smallest version of each module that can address all of the relevant policy issues, and remove any questions in that module that are not needed to analyze these policy issues. If the module is still too long, questions needed only to address the least important policy issues are deleted. This shorter module is checked again to see if it exceeds the provisional size limit. The general principle is that the most important policy questions are addressed first and additional issues are added until the module has reached the length that survey designers, in consultation with high-level policymakers, have set for it.

Third, after this has been done for all modules, survey designers should prepare a list of issues they think can be covered by the survey and give this list to the high-level policymakers, who will decide whether they would like to change the amount of space allocated to each module. The survey designers should tell the policymakers about the tradeoffs involved, working with them to ensure that the issues policymakers deem most important are addressed.

Ultimately, this process produces a list of modules to be included in the survey, the proposed length of each module to be included, and the specific objectives for the modules. This completes the second step of survey design. This step may need to be revisited later if results of the field test show that the questionnaires are too long or that there is room to expand the questionnaire.

## Notes

The authors would like to express their gratitude to Jere Behrman, Lawrence Haddad, Courtney Harold, John Hoddinott, Alberto Martini, Raylynn Oliver, Kinnon Scott, and Salman Zaidi for comments on an earlier draft.

1. This book is designed to provide a thorough review of international experience. However, new experience and knowledge will continue to accumulate after the book has been published. Therefore, until a new book is written, any new international-level information is probably most easily obtained from international researchers.

2. If geographic areas—rather than households—are the unit of observation, it may be possible to merge data from different surveys. However, this high level of aggregation yields less precise

results, raises issues of aggregation bias, and generally requires surveys with very large sample sizes.

3. A variety of external sources have been used to fund past LSMS surveys. World Bank loans have partially financed several LSMS surveys. Grants from various bilateral development agencies (especially from the United States, Scandinavian countries, and Japan) and multilateral development agencies (particularly the United Nations Development Programme and the United Nations Children's Fund) have wholly or partially financed a large share of LSMS surveys. In a few cases, grants from the World Bank research budget have supported LSMS surveys. Similar surveys, such as the World Bank's SDA surveys, RAND's Family Life Surveys, and a few other surveys in Africa, all receive a large share of their funding from external sources.

4. Most previous LSMS surveys have used two-stage sample designs. If a three-stage sample design is used, ID codes will be needed that identify both the primary and secondary sampling units of each household. An analogous comment applies to surveys that use four or more stages in their sample designs.

5. In large countries with federal systems, surveys can be performed for individual states. Such surveys usually have the same general purposes as national surveys, and have samples that are representative of the whole state.

6. Each module in the household questionnaire should also be included in the community questionnaire. See Chapter 13 for further discussion of the community questionnaire.

7. While all versions of the household enterprise module collect income information, only the standard and expanded versions of the agriculture module collect sufficient data for use in the measurement of total income.

8. A full LSMS-type survey could accommodate two and possibly three expanded versions of modules; a scaled-down survey could accommodate at most one. Volume 3 presents expanded versions of the following modules: roster, education, health, employment, migration, environment, household enterprises, and agriculture. The time-use modules introduced in Chapter 22 should also be treated as expanded modules, and the same is even more true for the full set of environmental modules introduced in Chapter 14.

## References

- Grosh, Margaret, and Juan Muñoz. 1996. *A Manual for Planning and Implementing the Living Standard Measurement Study Survey*. Living Standards Measurement Study Working Paper 126. Washington, D.C.: World Bank.



# 3 Designing Modules and Assembling Them into Survey Questionnaires

*Margaret Grosh, Paul Glewwe, and Juan Muñoz*

Chapter 2 outlined the five-step process that survey designers should follow to design LSMS and similar multitopic surveys. It also provided detailed recommendations on how to undertake the first two steps, which are deciding on the overall design of the survey and deciding which modules to include in the survey questionnaire. This chapter discusses the last three steps of the five-step survey design process. The first section of this chapter describes the third step—drafting each module, question by question, to ensure that it will collect the data necessary to meet the module’s objectives (which were laid out in the second step). The second section guides survey designers through the fourth step—coordinating the different modules and combining them to create a consistent and comprehensive set of questionnaires. The third section explains the procedures for the last step—translating the questionnaires into local languages and conducting a field test. The fourth section discusses the formatting of the questionnaires, which is an extremely important but often neglected aspect of designing successful multitopic surveys. Survey designers should refer to the material contained in the fourth section many times during the last three steps of the survey design process.

In practice, the survey design process rarely moves smoothly and sequentially from one step to the next. Instead, survey designers often find themselves moving backward and forward among the various steps. For example, if designers encounter difficulties when drafting a specific module, they may need to reconsider and modify their original objectives for that module. Developing survey questionnaires is an iterative process, and survey designers should expect to go through at least three or four drafts of each module. It is not unusual for the different versions of the drafts to add up to a stack of paper one foot (30 centimeters) high. Each major redraft of a module or questionnaire should be reviewed by all interested parties, not only

the people involved in carrying out the fieldwork (the data producers) but also policymakers (who will make decisions based on the data), members of the research community (who will analyze the data), and the staff of any agencies financing or providing technical assistance to the survey. Eventually, what should emerge from the process is a well-designed set of questionnaires for a multitopic household survey.

## **Producing Draft Modules**

The third step in survey design, producing draft modules for the household and community questionnaires, is one of the most time-consuming steps in the

process. Detailed guidance on this step is provided in the chapters in Parts 2 and 3, so the discussion here will be general and relatively brief.

Once the objectives for each module are finalized (at least tentatively), survey designers can begin to develop detailed draft modules for the household and community questionnaires. Survey designers should use the draft “prototype” modules introduced by the chapters in Parts 2 and 3 (and presented in Volume 3) as their starting point. As explained in Chapter 2, survey designers will already have decided on the policy and analytical objectives of each module. They should now choose the shortest versions of the modules that will allow for analysis of the most important of these objectives; any questions not relevant to these objectives should be removed.

If the resulting module is still too long, survey designers should remove any questions that are needed only for the analysis of the least important of the policy issues. This process should continue until the module meets the length constraint. In some cases the module may be shorter than expected, in which case a policy issue and its accompanying questions can be added. The general principle is that the most important policy issues should be addressed first, and additional ones should be included only if space allows. This approach is a good start, but much more remains to be done.

For some modules the information and guidance given in the relevant chapter in Parts 2 and 3 of the book may be incomplete. For example, the chapter may not address certain policy issues that are important in a given country or setting, in which case the designers of that survey will need to develop new modules or submodules. Even in these cases the information provided in the relevant chapter is usually a good base for developing such modules. However, if the designers intend to implement major innovations in their survey, they should seriously consider adding to the survey team a specialist with the relevant experience in both data collection and data analysis.

Once each draft module has been written out in its entirety, the next task is to verify that the design of each module reflects the economic and institutional structures of the country in question. For example, the designers need to check whether common living arrangements are reflected in the definition of the household used in the household roster and in the housing and interhousehold transfer modules. They

must also review all questions and response codes and, if necessary, modify them to reflect local institutions and terminology. For example, the transfers and other nonlabor income module discussed in Chapter 11 must explicitly refer to each public transfer program by name. The consumption module will need even more work; in particular, the lists of items selected must closely reflect items consumed in the country. The agricultural module will need careful attention, as this module must reflect the country’s landholding and cropping patterns.

For many of the modules, survey designers may find it useful to collect some preliminary data using qualitative techniques, which may help them determine how best to design these modules to collect quantitative data. Chapter 25 provides a detailed discussion on how to collect qualitative data. Such data can be particularly useful in countries where successful quantitative surveys have never been done for the topic to be studied.

A final general issue to consider when drafting modules is the role played by the fieldwork schedule. A prototypical full LSMS survey spreads fieldwork evenly over a 12-month period, for two reasons. First, this makes it possible to study or average out any seasonality effects. Second, and more importantly, surveys with this fieldwork schedule require a smaller number of survey field teams than do surveys that compress the fieldwork into a shorter period of time. This smaller number of teams reduces costs and allows for improved quality control. All of the interviewers can be trained together and thus to a uniform standard; in addition, the cost of training interviewers—which takes about four weeks—will be proportionately cheaper. Each interviewer will complete more interviews and thereby gain more experience. Finally, fewer computers and vehicles will be required.

Despite these advantages of a year-long survey period, many past LSMS surveys have compressed fieldwork into a period of just two or three months. This has often been done when there was pressure on the survey team to collect data for analysis as quickly as possible. In other cases interviewers may have been available for only a short period of time, or the organization funding the survey may have required that the project be completed in a relatively short amount of time. The fieldwork schedule can also be modified to accommodate analysis of certain topics. For example, analysis of some agricultural issues may require inter-

viewers to make two or more visits to each household at different times during the year.

Variations in the fieldwork plan may require changing the wording of some modules. This means that survey designers should ensure that the design of each module in the questionnaire is consistent with the fieldwork plan.

When a survey is conducted over a relatively short period, such as a few weeks or months, careful attention must be given to the wording of questions concerning events that are seasonal in nature. Will school be out of session for a large portion of the survey period? If so, the education module may need to be changed to reflect this. In particular, questions referring to school activities during the previous week, such as the number of days that a child attended school or the number of hours of homework done by the child, would clearly be inapplicable. Also, questions about water supply during both wet and dry seasons should be reviewed to ensure that they reflect the circumstances of these seasons. The largest seasonal changes may need to be made to the agricultural module. A detailed discussion of the implications of seasonality for that module is provided in Chapter 19.

More substantial changes will be required if the household is to be visited more than once at different times of the year. In such cases it may be desirable to have the interviewer administer modules for which the answers are expected to vary by season (such as the consumption, agriculture, water, or time use modules) each time he or she visits the household. In contrast, the modules that are unlikely to be affected by seasonality, such as housing, education, fertility, or migration, probably need to be administered only once. Any modules that are to be administered more than once usually need to be modified, particularly with respect to their recall periods. For example, if the interviewer makes two household visits six months apart, the consumption module should be administered in both visits and should have a recall period of six months rather than one year. Also, the water module should ask only about the particular season (wet or dry) during which the interview is to be conducted.

The guidelines given in this chapter are general, since very detailed information is provided in Parts 2 and 3 of this book. Other information on adapting LSMS questionnaires to fit local circumstances can be found in Oliver (1997), which focuses on survey design in the countries of the former Soviet Union, as

well as in Ainsworth and van der Gaag (1988). A good general reference publication for developing and designing household survey questionnaires is United Nations (1985). More recent general references are in Babbie (1990), Fink (1995), and Fowler (1993); although these books focus more on developed countries, much of the material they contain is also relevant for developing countries. A final point to bear in mind is that a good deal of attention must be given to correct and consistent formatting. This is described in great detail in the fourth section of this chapter; survey designers should read that section very carefully before they begin designing any survey modules.

### **Integrating and Combining Modules to Create Complete Questionnaires**

Once draft versions of each of the individual modules have been written, these drafts must be combined to form complete household and community questionnaires. Merely stapling the various modules together will not produce a well-designed questionnaire; much more work has to be done to ensure that the different modules fit well together. This section describes how to do this important task. It focuses primarily on making the modules of the household questionnaire consistent with each other. Similar, though less difficult, issues arise when integrating the modules of the community questionnaire; in most cases the approach to take for the community questionnaire can be inferred from the discussion of the household questionnaire. This section will also highlight particularly important points to consider when combining the household, community, and price questionnaires to form a comprehensive household survey.

#### **Gaps and Overlaps**

Survey designers must scrutinize and compare the different questionnaire modules for gaps and overlaps in the information that the modules collect. Analysts often need to combine data from different modules in the household questionnaire. Perhaps the most important example of this is the calculation of each household's total consumption, which requires information not only from the consumption module but also from the education, health, employment, and housing modules—and from the water, sanitation, or fuel modules (see Chapter 14) if they are included as separate modules in the questionnaire (as opposed to using the

housing module to collect information on these topics). Likewise, income data are collected in the employment, agriculture, household enterprise, and miscellaneous income modules. It is important to check that a questionnaire includes the data needed to construct these and other complex variables.

Another example of this general issue is that survey designers often have a choice regarding the module in which to collect some kinds of information. For example, data on expenditures on fuel for cooking and heating can be collected in the consumption module, the housing module or, if it exists, the expanded fuel module. Questions on child immunization can be placed in the fertility module, the health module, or the anthropometry module. An argument can be made for choosing any of these options (see the pertinent chapters in Parts 2 and 3 of this book), but the essential point is to ensure that the information is collected at least once, and is collected twice only if there is a reason to do so.<sup>1</sup> Appendix 3.1 provides a list of the most common types of gaps and overlaps to check.

In cases in which information could be plausibly be collected in more than one part of the questionnaire, there may be no absolute right or wrong place to collect it. Rather, survey planners must take into account who the respondent is in each module, how well the best recall period for that information matches the recall periods of modules in which it might be collected, at what point in the interview the respondent might discuss the topic most naturally, and whether the topic is a sensitive one that should therefore be addressed near the end of the questionnaire (for reasons discussed further below).

The survey designers should also examine any overlaps among the household, community, and price questionnaires. In general, the community and price questionnaires should collect information on any topic that varies only slightly from household to household within the primary sampling unit.<sup>2</sup> While much of the information collected in the community and price questionnaires could be collected in the household questionnaire, it is better to collect it in the community questionnaire in order to shorten the length of each household's interview. Collecting this information in the community questionnaire is also more efficient; why collect it for all households in a primary sampling unit (often 16 or 20 households) when it need be asked only once in the community questionnaire?

Some simple examples illustrate this point. The expanded water module contains questions about the price and quality of water from different potential water sources. If the primary sampling units are geographically compact, all of the households in each primary sampling unit are likely to have the same alternative water sources, implying that the water price and quality questions can be put in the community questionnaire (which should be administered in each primary sampling unit) rather than in the household questionnaire. On the other hand, if the primary sampling unit is not compact so that the households are widely dispersed, it is likely that some households will be nearest to, say, a particular spring or well while other households will be closer to other springs or wells. In such cases these questions about alternative water sources should remain in the household questionnaire.

Another example concerns the distance to schools and health facilities. In a compact primary sampling unit, the distance to the nearest school or health facility probably varies little among the households in the primary sampling unit. This means that information on the distances to schools and health facilities can be collected in the community questionnaire as opposed to the household questionnaire.

### *Length*

The overall length of the household questionnaire must be manageable. In general, it is not feasible to include, say, the standard version of each module presented in Volume 3, even though past full LSMS surveys typically included 15 modules, many of which were similar to the standard versions in this book.

There are several reasons why using all of the standard draft modules in this book is not feasible. First, this book introduces several new modules, including the time use module and several environmental modules. Second, some of the standard draft modules, such as those on health, migration, and household enterprises, are much longer than the modules on those topics that were used in previous LSMS surveys. Finally, in some of the chapters in this book (including Chapter 18 on household enterprises and Chapter 19 on agriculture) it is argued that collecting more detailed data will greatly increase their value for analytical purposes. Thus survey designers should not combine the standard versions of all of the modules presented in the book into a single household questionnaire. Instead, the short versions should be used for

some modules, and in almost all cases at least one or two modules should be dropped.

Assessing whether a draft questionnaire is too long is not simply a matter of counting the pages or questions in it, since many questions, and sometimes even entire pages or modules, will apply only to some households. Moreover, in some cases adding questions does not lengthen the interview time because the respondent cannot avoid going through the thought process made explicit in these questions, which implies that a supposedly abbreviated set of questions will not reduce the time required to complete the interview. An example of this is the calculation of income derived from agricultural activities.

There are also several ways to implement long questionnaires that minimize the time required by (and the fatigue induced in) each survey respondent. These include conducting individual “mini-interviews” with each household member to collect all of the information needed from that individual at one time (which allows him or her to leave when questions are being asked of other household members); using the best-informed respondent for each household module; and dividing the interview into multiple visits (for example, going through all the individual-specific modules in one visit and returning on a different day to conduct the consumption module and other household-level modules). LSMS surveys use all of these techniques. Still, there is a limit to the amount of information that can be gathered from a single household.

How can survey designers determine whether a household questionnaire is too long? A rough idea of the effective length of the questionnaire in different circumstances can be obtained by calculating how many households will go through the different paths created by the skip patterns and how many questions will be asked for each possible path. An excellent example of this is provided in Chapter 18 on household enterprises, in Table 18.5.

A more precise estimate of the time required to administer a household questionnaire can be obtained when similar surveys have already been done in the country or region studied. In this case, the designers of the new survey will be able to find out how long the interviews took in the previous survey, provided that the earlier survey collected metadata along the lines suggested in Chapter 4. If such information is not available, survey designers will need to rely on the field

test, which is discussed in detail in the fourth section of this chapter. If field test interviews require many hours to complete and exhaust the cooperation and patience of households, this is an indication that the questionnaire is too long. At the same time, survey designers should realize that field test interviews normally require much more time than do similar interviews during an actual survey, because interviewers have little training or experience with the questionnaire at the time of the field test. In addition, the questionnaire used in the field test is not a final draft and thus is likely to contain some problems that will slow down the interviews. A handy rule of thumb is that interviews in the actual survey take only about half of the time that they take in the field test—and sometimes even less than that.

A general goal to aim for in the actual survey is that any given respondent should not be interviewed for more than one hour on a given day. Of course, people’s tolerance for being interviewed will vary from country to country, and this general guideline must be adapted to suit local conditions. Experience in LSMS surveys to date suggests that people’s tolerance for long interviews is lower in urban areas than in rural areas, lower among wealthy households than among poor ones, and lower in wealthier countries than in poorer ones.

#### *Recall Periods*

The recall periods proposed for each module introduced in this book are mostly those that the authors have deemed appropriate for that particular module. This can be a problem when analysts want to combine or compare data from several modules. For example, in many LSMS surveys the employment module uses a one-week recall period. Since most adults work, this yields a large number of observations, and the period of time is short enough to yield accurate answers to such basic questions as the number of hours worked and the payments received during this recall period. In contrast, the health module uses a four-week recall period. This relatively long recall period is used because most people are not ill in any given week. The four-week recall period allows more observations of illness for a given sample size than would be obtained using a one-week recall period. Since illnesses are important events, respondents can be expected to remember many details of their episodes of illness during the past four weeks.

However, if an analyst wants to study the impact of illness on earnings or work effort, these different recall periods will complicate the analysis. The analyst cannot tell whether the illness took place before or during the period for which the earnings and hours data were collected. This could be resolved either by adding questions to the health module to specify the days during the recall period on which the respondent was ill or by making the recall periods coincide, perhaps with a compromise of two weeks for both modules (bearing in mind the disadvantages in sector-specific analyses of using a recall period different from the “ideal” one for that module). Part of the job of integrating the draft modules is to determine and judge the tradeoffs being made, either confirming that they are acceptable or altering them until a more appealing tradeoff is reached.

#### *Nomenclature and Coding Schemes*

The questionnaire should be reviewed to check that wherever similar questions are asked, the nomenclature and coding schemes are the same. This should reduce coding errors and simplify data analysis. For example, many different modules allow the respondent to choose the time unit (for example, hour, day, week, or month) that they find most convenient when responding to questions regarding time or payments over time (such as wage rates, the length of time spent gathering firewood, and the length of time covered by a payment for water). The code numbers for these time units should be the same throughout the entire questionnaire; in the draft modules presented in this book “day” is always coded as “3,” “week” is always coded as “4,” and so on for other units of time.

Another example concerns the migration module, the transfers received page of the transfers and other nonlabor income module, and the transfer payments page of the consumption module. All have questions about where the migrant, donor, or recipient lives. The coding scheme that categorizes this information, whether it is the type of place (capital city, other urban area, rural area, or overseas) or the name of the place, should be uniform. Likewise, several modules include questions about the relationship between two individuals. It is usually a good idea for these questions to use the same codes that are used in the household roster module to indicate the relationship of each household member to the head of the household.

A particularly important task is to coordinate the coding of items in the consumption expenditure module with items in the price questionnaire. As explained in Chapter 2, price data are needed to generate regional and temporal price indices that enable comparison of real expenditures of households interviewed in different places and at different times. This is done by matching the prices collected in the price questionnaire with the consumption expenditure information gathered in the consumption module. If the items are not well matched, this task becomes more difficult, and the resulting price indices will be less accurate. In general, the goal should be a one-to-one correspondence between the items listed in the consumption module and the prices collected in the price questionnaire. For example, if questions are asked on two or three varieties of rice or wheat in the consumption module, a price for each variety should be collected in the price questionnaire.

This should be relatively simple to do for almost all food items. Nonfood items are more difficult. It is usually not possible to obtain prices for durable goods because they often come in many varieties (for example, there are many kinds of bicycles or televisions). However, for nondurable items, prices can be obtained for well-defined examples. For example, there are many kinds of shirts, but if a specific widely purchased type of shirt can be defined, data on that type of shirt can be collected in the price questionnaire and used as an indicator of prices for all kinds of shirts. See Chapter 13 for a detailed discussion of the price questionnaire, including a list of suggested food and nonfood items to include in it.

#### *Choosing the Order of the Modules in the Household Questionnaire*

A final and very important question to address is the order of the modules in the household questionnaire.<sup>3</sup> It is natural and convenient to arrange the modules in the order that they will be administered, so the key issue here is the order in which the modules will be administered and how this affects the physical design of the questionnaire.

To put this issue in context, consider the traditional fieldwork plan for a full LSMS survey. Each field team works in its assigned primary sampling units (communities) twice. The first time a team arrives in a primary sampling unit, it works there for about one week. The first half of the questionnaire, most of

which usually consists of the individual-specific modules, is completed for each household. In addition, a short module is administered that asks which household members are best able to answer questions concerning the specific household-level modules (agriculture, household enterprises, consumption, and savings) that will be filled out when the team returns to the primary sampling unit about two weeks later. Figure 3.1 provides an example of such a module.

The field team works in a different primary sampling unit during the following week, while the data in the half-completed questionnaires from the first primary sampling unit are entered into a computer by a data entry operator (who does not travel with the team) using a data entry program. The data entry program checks the first half of the questionnaires for a wide range of errors and inconsistencies. (This is discussed more fully in Grosh and Munoz 1996.) The team returns to the first primary sampling unit in the third week, administers the rest of the questionnaire (which mainly consists of household-level modules), and resolves any problems or inconsistencies found by the data entry program when the data from the first half of the questionnaire were entered.

In several recent LSMS surveys, two different procedures have been used in the fieldwork stage. One procedure is that the data entry operator travels with the field team. This option has become feasible with the advent of small laptop computers that can be powered by batteries, vehicle cigarette lighters, or solar panels. This allows the whole questionnaire to be filled out and checked using the data entry program during a single trip to the primary sampling unit. In addition, the second half of the questionnaire can be checked by the data entry program almost immediately, so that interviewers can return to the sampled households to resolve any problems detected by the program.

The other procedure, used when a scaled-down LSMS survey is being implemented, is to complete all of the interviews in a single trip to the primary sampling unit and sometimes even in single visits to each household. This procedure will have a serious disadvantage if the data entry operator does not travel with the team, because none of the data can be checked in time to return to the households to resolve problems detected by the data entry program. If the data entry operator travels with the team, there is little difference between this procedure and the former procedure, except that a full LSMS questionnaire will require

more visits to each household during the sole trip to the primary sampling unit.

Given these different possible fieldwork plans, there are several basic principles about how to order the modules in the household questionnaire. The first principle is that any modules on topics that respondents might consider sensitive should be put at the end of the questionnaire. This gives the interviewer time to develop a rapport with the household members, which should increase the probability that they will answer questions on sensitive issues, and do so truthfully. It also means that if the respondent breaks off the interview in response to a sensitive question, only the data from that last module or modules are lost. Finally, by this point in the interview, any interested onlookers, such as family members and neighbors, may have wandered away, making it possible to administer the more sensitive portions of the questionnaire with greater privacy. Education, housing, migration, and in some cases health<sup>4</sup> are usually good topics with which to open the interview, because people generally do not mind talking about these topics. In contrast, fertility, savings, credit, and transfers and other nonlabor income are among the most sensitive topics in the household questionnaire.

A second principle concerns bounded recall periods. In past LSMS surveys in which the interviewer made two visits two weeks apart to each household, some parts of the questionnaire used bounded recall periods; in other words, questions were asked such as "How much has your household spent on rice since my last visit?" As explained in Chapter 5, using bounded recall periods can increase the accuracy of the respondents' answers. Obviously, if bounded recall periods are used in certain modules, these modules must be administered in a second visit to the household and thus be included in the second half of the questionnaire. The two modules in Volume 3 that explicitly use bounded recall periods are those on consumption (Chapter 5) and household enterprises (Chapter 18).<sup>5</sup>

A third consideration is the selection of respondents. As explained above, several modules (including the consumption, agriculture, household enterprises, savings, housing, and environmental modules) collect much or all of their data at the household level, which means that the questions are answered by the household member most knowledgeable about that topic. With the exception of the housing module, these

FIGURE 3.1: MODULE FOR CHOOSING RESPONDENTS TO BE INTERVIEWED IN THE SECOND HALF OF THE QUESTIONNAIRE

RESPONDENTS FOR SECOND HALF OF QUESTIONNAIRE

RESPONDENT: THE PERSON BEST INFORMED OF THE ACTIVITIES OF THE HOUSEHOLD MEMBERS

FULL NAME OF THE RESPONDENT: \_\_\_\_\_ ID CODE

1. Who shops for the food for your household?

NAME: \_\_\_\_\_ ID CODE

2. Who in your household knows most about the non-food expenses of the members of your household?

NAME: \_\_\_\_\_ ID CODE

3. Who in your household knows most about the miscellaneous income and transfers received from other households?

NAME: \_\_\_\_\_ ID CODE

4. Who in your household knows most about the savings in your household?

NAME: \_\_\_\_\_ ID CODE

5. During the past 12 months has any member of your household participated in agricultural production, forestry, or raising livestock?

YES.....1   
NO.....2(»8)

6. Who is the person who knows most about all the agricultural and livestock activities of the members of your household?

NAME: \_\_\_\_\_ ID CODE

7. In addition to this person, who else in your household manages plots of land owned or rented in by the household? Who is responsible for plots that are rented out by the household?

NAME	ID CODE

**FIGURE 3.1: MODULE FOR CHOOSING RESPONDENTS TO BE INTERVIEWED IN THE SECOND HALF OF THE QUESTIONNAIRE**

<p>8.</p> <p>Over the past 12 months, has anyone in your household operated any non-agricultural enterprise that produces goods or services (for example, artisan, metalworking, tailoring, repair work, and processing and selling your outputs from your own crops, if done regularly) or has anyone in your household owned a shop or operated a trading business?</p> <p style="text-align: center;">YES . . 1 NO . . . 2 (»NEXT MODULE)</p> <div style="text-align: center; border: 1px solid black; width: 40px; height: 20px; margin: 0 auto;"></div>	<p>9.</p> <p>What kind of enterprises does your household operate?</p> <div style="border: 1px solid black; padding: 5px; text-align: center; margin: 10px auto; width: 80%;">                 PROBE TO DETERMINE INDUSTRIAL SECTOR IN WHICH ENTERPRISE OPERATES.             </div>	<p>10.</p> <p>Who is most informed about and/or in charge of day-to-day operations of the enterprise?</p>																														
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 10%;">ENTERPRISE ID</th> <th style="width: 45%;">FULL WRITTEN DESCRIPTION</th> <th style="width: 10%;">CODE</th> <th style="width: 25%;">NAME</th> <th style="width: 10%;">ID CODE</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">1</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td style="text-align: center;">2</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td style="text-align: center;">3</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td style="text-align: center;">4</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td style="text-align: center;">5</td> <td></td> <td></td> <td></td> <td></td> </tr> </tbody> </table>			ENTERPRISE ID	FULL WRITTEN DESCRIPTION	CODE	NAME	ID CODE	1					2					3					4					5				
ENTERPRISE ID	FULL WRITTEN DESCRIPTION	CODE	NAME	ID CODE																												
1																																
2																																
3																																
4																																
5																																

modules are quite lengthy. Thus for each of these modules it is usually best for the interviewer to ask which member would be the most appropriate respondent during the first visit to the household, and then make an appointment to interview that person at a later, more convenient time. In the traditional two-visit fieldwork plan, this implies that these modules, except perhaps housing, should be administered during the second visit and thus should be located in the second half of the questionnaire. However, if the team travels only once to the primary sampling unit, it is still feasible to make appointments for later in the day or for another day, which gives survey designers more flexibility in deciding where to place these modules in the questionnaire.

A fourth principle relates to the logistics of data entry. The individual-level modules include many more questions for which strict range and consistency checks can be built into the data entry program than do the modules on consumption, agriculture, and household enterprises.<sup>6</sup> If the whole questionnaire is completed using the traditional LSMS fieldwork plan (two visits two weeks apart), all the individual-specific modules except the credit module should be administered during the first interview. (The credit module is probably too sensitive to be administered during the first interview.) This will allow the survey team to enter the data from these modules and to detect any apparent errors or inconsistencies that could then be resolved in the second interview. If the data entry operator travels around with the interview team, the data from the interviews can be checked in a matter of hours; thus where these modules appear in the order of the household questionnaire becomes less important for the purposes of data entry.

Given these principles, and some common sense, more specific advice can be given. Each household questionnaire should have the metadata module at the very front, since much of the information that module collects (such as whether the interviewer successfully located the household, the date of interview, and the language in which the interview was conducted) becomes apparent at the very beginning of the interview. The next module should be the household roster; this must be completed before any other module because it determines who is and who is not a household member, and thus determines the people to whom all the other modules will apply. However, at least one page of the household roster, the one with

the names of all the household members, is usually placed further back in the questionnaire so that the names on that page can be seen during the administration of all individual-level modules. Thus the physical placement of this page will not reflect the time during the interview when it is filled in. (For details see the discussion on the fold-out roster page in the fourth section of this chapter.)

After the household roster, it is useful to fill out the form on selecting household respondents shown in Figure 3.1; this form can be administered to the same person who answered the household roster questions (usually the head of household or the person most knowledgeable about other household members). It is useful to collect this information early because it can be used to save household members' time by interviewing them sequentially using "mini-interviews." That is, after the interviewer has administered the form that identifies the relevant respondents for the household-level modules, he or she should administer all of the modules that are clearly individual-specific (except the credit and fertility modules) to each household member, finishing all such modules with one member before interviewing another member. These are the education, health, employment, migration, and time use modules. Some household members will not need to be interviewed further, and thus their mini-interviews will consist of the interviewer administering only these modules. In contrast, other household members will also be the respondents for some of the household-level modules. For example, the respondent for the housing module can also provide answers for the questions in that module as part of his or her mini-interview. Using this method, the interviewer can obtain all of the information needed from each individual in a way that minimizes the use of respondents' time; once a respondent finishes the mini-interview he or she can leave or start some activity without further interruption.

Within this group of individual-level modules, those on education and migration should be administered first since the information they collect is not very sensitive. Some employment information can be sensitive, particularly questions concerning wages, so this should be one of the last of the individual-level modules to be completed, if not the very last. If the short health module is used, it can be put near the front. However, if the standard or expanded version is used, it should be placed toward the end of the individual-

specific modules because of the sensitive nature of some of the questions in this version (see endnote 4).

Which modules should go near the end of the questionnaire? Because the three most sensitive modules are those that collect information on savings, credit, and transfers and other nonlabor income, these three modules should probably be put at the end of the questionnaire. Another potentially sensitive topic is fertility. In countries in which fertility is particularly sensitive, it should come immediately before the savings, credit, and transfers and other nonlabor income modules.

Where should the other modules go? If the traditional “two visits two weeks apart” interview system is used, the consumption and household enterprise modules should be in the second half of the questionnaire since these modules often use a bounded recall period, namely the time since the interviewer’s previous visit. Modules that are long and also need to be administered to specific respondents—the consumption, agriculture, household enterprise, and environment modules—should also be completed in the second visit. Finally, as discussed above, the housing module can be administered in the first interview because it is unlikely to contain any sensitive questions. If the two-visit system is not used, these modules can be put anywhere between the individual-specific modules and the more sensitive modules.

Finally, the goal of saving respondents’ time by conducting mini-interviews with each respondent (who can leave after his or her mini-interview is finished) is complicated by the fact that the household enterprise, agriculture, miscellaneous income, and credit modules consist of a mixture of household-level and individual-level questions. For example, in the standard and expanded versions of the agriculture module, individual household members are asked whether they have worked on specific plots of land. However, these questions cannot be asked until several other questions have been asked about the different plots of land owned and rented in by the household members—and such questions would be awkward to ask in a form as simple as the one shown in Figure 3.1.

The best way to resolve this problem will depend on which modules and which versions of these modules are included in the household questionnaire, so it is difficult to provide general advice. However, one way to reduce the time burden on household members is to identify all of the people

who still need to be interviewed after the “mini-interviews” are completed, as this will allow people to leave if they do not need to be interviewed further. Continuing the agriculture example, note that the form in Figure 3.1 identifies all of the household members who either manage or work on a plot of land. Household members who do not fit this description and who are not needed to complete any other household-level module can leave after their “mini-interview” is finished.

This completes discussion of the fourth step of integrating the draft modules and combining them into a complete set of questionnaires. The primary focus has been on the household questionnaire, since the community questionnaire is much smaller. (See Chapter 13 for a detailed discussion of the community questionnaire.) Designers of prospective surveys can consult the questionnaires used in previous LSMS surveys by downloading them from the LSMS website, <http://worldbank.org/lms/lms/home.html>.

### **Translating and Field Testing the Draft Questionnaires**

After the draft modules have been combined into a complete set of household, community, and price questionnaires, they need to be translated and field tested.<sup>7</sup> The field test is particularly important because it is the last check on the design of the questionnaires before the survey is implemented.

#### *Translation*

It may be necessary to translate the questionnaires for three reasons, each of which has different implications for the design of the survey. The most common and most important reason is that respondents may speak a range of different languages. In many countries more than one language is spoken. In these countries quality control requires that a separate questionnaire be produced for each of the major languages spoken in the country, with every question written out verbatim.

Scott and others (1988) demonstrated how this procedure greatly increases the accuracy of the data collected. They conducted an experiment designed to measure interviewer errors when the interviewer had to translate each question during the interview. For example, the interviewer may have had to use a questionnaire written in English to conduct an interview in Tagalog or Cebuano or a questionnaire written in

French to conduct an interview in Baoule or Dioula. The interviewers' error rates were two to four times higher when they translated questions during the interview than when they used questionnaires already written in the languages used by the respondents.

While the final versions of the questionnaires must be translated from the national (official) language to produce verbatim questionnaires in the other languages used in the country, the preliminary drafts of the questionnaire can be developed using only the national language. Ideally, the version of the questionnaire to be used for the field test should be translated into each of the languages that will have a final written version of the questionnaire. In practice, field tests are often done using only oral translations of the national language version of the questionnaire. Thus the wording in the local language interviews during the field test may not correspond exactly to the wording that will be used in the written translations of the final questionnaire. While this is an imperfect way to proceed, it is often a reasonable tradeoff given the high costs, in both time and in money, of field testing the questionnaires in each language.

After the final version of the household questionnaire has been translated into another language, the translation needs to be carefully checked. The best way to do this is to use "back translation." That is, after the questionnaire is translated from the language in which it was developed into the languages in which it will be administered, someone should translate the versions in those languages back into the original language. After this "back translation" has been accomplished, the two versions in the first language should be checked. Where there is a discrepancy in wording or meaning between the two versions, the translation should be carefully checked. A person or group of people familiar with the purpose of the questions should do the first translation. The back translation should be done by someone who was not intimately involved in designing the questionnaire. Any ambiguities and errors must be noted and corrected in the translated version rather than being "fixed" only in the back translation version.

Most previous LSMS questionnaires were printed only in the national languages of the countries studied, so multilingual interviewers had to be employed to conduct interviews in the most commonly used local languages. Occasionally a few key questions or phrases were translated into the local languages and

written down in the interviewers' manual. In the case of the least common languages, local interpreters were used when none of the interviewers spoke the language. In this respect, while previous LSMS surveys have conformed to normal survey practice, they have not reached the cutting edge of quality control as defined by the World Fertility Surveys. The guidelines used in those surveys require that questionnaires be prepared in all languages used by more than 10 percent of the sample and that a minimum of 80 percent of the sample be interviewed using questionnaires written in the respondents' native language.

Future LSMS and similar multitopic surveys should make greater efforts to translate the household questionnaire into local languages. When preparing these translations, the questionnaire should always be worded in the way that the language is commonly spoken, using relatively simple terms and avoiding academic or formal language. The gap between the spoken and written languages and the difficulty of striking a balance between simplicity and precision may be greater in local languages, especially ones that are not commonly used for reading and writing. The translators should therefore be especially careful to try to find an appropriate balance.

Two examples illustrate the kind of problems that can occur. The question "¿Estuvo enferma en las últimas cuatro semanas?" literally asks, in Spanish, whether the respondent was sick in the past four weeks. However, in spoken Spanish in Chile it could be understood as a polite euphemism for asking whether a woman has had a menstrual period in the last four weeks. An even more difficult problem in wording was revealed in the field test in Nepal. Apparently the most natural Nepali phrasing for "Have you been ill?" is closer to "Have you been to the doctor?" The change in meaning from what was intended appeared in the field test several times when respondents answered "No, I couldn't afford to go," clearly an inappropriate response to the question "Have you been ill?"

The second reason why the questionnaires may need to be translated is that sometimes the international experts working on the survey design team do not speak the national language well enough to design the questionnaires in that language. This happened in the case of the Vietnam LSMS questionnaires, which were developed jointly in English and Vietnamese. In contrast, the LSMS questionnaires used in Latin American countries have been drafted only in Spanish by teams of

local and international experts who are fluent in that language. When translation is a necessary part of the development of the questionnaires, each draft of the questionnaire must be translated, which may require a substantial amount of money and can also increase the time needed for designing the questionnaires.

The third and final reason for translating questionnaires is to produce a questionnaire in one or more of the major international languages (English, Spanish, or French) in order to encourage the international research community to use these data in their policy analysis. Such translations need not be done until after the final questionnaire is developed, and back translations are not needed.

### *Field Testing*

After draft versions of the household, community, and price questionnaires have been assembled and (if necessary) translated, they must be tested in the field. The field test is one of the most critical steps of the survey design process. The goal is to ensure that the questionnaires are capable of collecting the information that they are intended to collect. A field test should address the adequacy of the draft questionnaires at three levels:

- *The Questionnaire as a Whole.* Is the full range of required information collected? Is the information collected in different parts of the questionnaire consistent? Are any variables unintentionally double-counted?
- *Individual Modules.* Does the module collect the intended information? Have all major activities been accounted for? Are all major living arrangements, agricultural activities, and sources of in-kind and cash income accounted for? Are some questions missing? Are some questions redundant or irrelevant?
- *Individual Questions.* Is the wording clear? Do any questions allow for ambiguous responses? Are there multiple interpretations? Have all responses been anticipated and coded?

It is important for a field test to include households from all major socioeconomic groups. For example, a sample should include: rural and urban households; individuals employed in the formal sector, in the informal sector, and in agriculture; and farmers in each main agroecological region, in each production scheme (independent farming, renting, sharecropping, and cooperative farming), and so forth. The

households should not be selected at random. Instead, different types should intentionally be included so that all of the various situations likely to be found during the survey are observed during the field test.

Experience with LSMS surveys has shown that field tests should be conducted using at least 100 households. To get enough responses for each module of the questionnaire, it may be necessary to visit additional households to conduct partial interviews in which only those modules that apply to a relatively small number of households are administered. For example, the original 100 households may not include enough pregnant women or people who have been ill in the month preceding the interview to determine whether the fertility and health modules, respectively, are well designed. In such cases survey designers should find additional households that contain pregnant women or ill people and have interviewers administer only the fertility or health module to those households.<sup>8</sup> A field test usually takes about one month to complete—about one week for interviewer training, two to three weeks of fieldwork (interviewing), and one or two weeks to discuss the findings and finalize the questionnaires. More time is required if the final questionnaires are to be produced in more than one language, because each version of the questionnaire should be field tested.

While the full field test should cover 100 or more households, much can also be learned from preliminary smaller tests. A general rule of thumb is that about half of the problems will show up in the first 10 households interviewed. In one recent field test, international experts wrote six pages of comments about a single module after interviews were completed for only three households. Such small-scale preliminary field tests are often particularly appropriate for new or difficult modules. Yet survey designers must understand that these are precursors to a full-size field test of the whole questionnaire, and not a substitute.

The personnel involved in a field test should include the survey design team, a few experienced interviewers or field supervisors, and a few of the people consulted by the survey design team, including both policymakers and research analysts. It may also be helpful to include people with experience working on past LSMS or similar multitopic surveys. All of the participants should divide into a small number of teams, each of which includes at least one person with each kind of expertise.

There should only be a few teams involved in the field test, usually around three or four. Mechanisms should be set up to enable the teams to contact each other during the field test so that they can compare notes on the problems they encounter and the solutions they have tried. A good way to set up such mechanisms is to have all of the teams working together for the first few days, perhaps in the capital city. This means that the teams will be in contact with each other every evening during the period when the first and often biggest flaws in the draft questionnaire are uncovered. In some cases the team members can agree on modifications to the questionnaire during the field test itself, which allows these modifications to be field tested.

Each interview during the field test should include, at minimum, the respondent, the interviewer, and an analyst or senior survey specialist. During the field test it is acceptable for the analyst or survey specialist to interrupt the interview tactfully in order to refine the wording of a question or the responses coded for it. Of course, in the actual survey the interviews should be conducted in private, and the interviewers should adhere to the wording of the questionnaire.

The interviewers used in the field test should be drawn from the experienced staff of the statistical agency. They should be good interviewers—familiar with basic interviewing practices and able to distinguish between problems caused by deficiencies in the questionnaire and problems caused by their lack of familiarity with the questionnaire. The interviewers' training should focus on the purpose of the survey and the structure and format of the questionnaire. One week of training is usually sufficient, followed by two or three weeks of household interviews.

Survey planners should set aside 1–2 weeks immediately after the field test to review the field test results and debate how to modify the questionnaire in light of those results. The group involved in the field test should go through the questionnaires, module by module, and discuss any problems that arose. At this stage, the team should bear in mind that the length of time required for each interview will fall dramatically when the interviewers are well trained and have become familiar with the questionnaire. As mentioned above, the typical field test interview will be at least twice as long as the average interview in the actual survey.

The data from the field test should not be entered in the computer or examined for any analytical pur-

poses, because in most cases the sample is both non-random and very small. However, the questionnaires from the field test can be used to check the performance of the data entry program.

The personal participation of all senior staff (including analysts) is fundamental for both the field test and its evaluation. The following anecdote illustrates this point. In one country, before the field test a manager in the statistics office asserted that collecting information on family assets would be impossible because respondents would fear that the information would be used for tax purposes. The module was included in the field test, and no unusual difficulties were encountered. But the manager who opposed the module did not witness the field test, and some of those who did participate in the field test did not participate in the module's evaluation. Despite the successful field experience, the module was removed from the questionnaire, largely because key decisionmakers did not fully participate in the survey design process.

Many small changes are generally made to questionnaires as a result of field testing, including changes in the wording of some questions, in questionnaire format, and in answer codes. If either the questionnaire's structure or the way in which certain variables are measured is changed substantially, all of the parts of the questionnaire that have been so modified must be tested again. This can delay the survey, but one way to reduce the probability of such a delay is to begin the field test with two or more versions of the most difficult, contentious, or important modules in the questionnaire. If one version clearly works the best, there is no need to do another field test because that version has already been field tested.

Ideally, the household, community, and price questionnaires should all be field tested at the same time. This allows the survey design team to evaluate all of the questionnaires together, taking into account the possibility that changes in one questionnaire may have implications for the design of the others. Simultaneous testing of the three questionnaires can also reduce travel costs since, like the household questionnaires, the community and price questionnaires should be tested in a variety of locations.

Regrettably, in several past LSMS surveys the survey teams neglected to field test the community and price questionnaires, concentrating solely on the household questionnaire. The community and price modules were tested late and haphazardly or, in some cases, not

tested at all. It is probably not coincidental that the users of the data from many previous LSMS surveys have often had more complaints about the community and price data than about the household data. If there is not enough staff time to test all three questionnaires at once, it is important to ensure that separate, rigorous field tests are done of the community and price questionnaires.

The health and education modules discussed in this chapter often include detailed facility questionnaires (in other words, school or health clinic questionnaires), which can be very complex (see Chapters 7 and 8 for details). It is essential to field test these facility questionnaires. During the field test the survey team should be sure to visit each type of facility covered by the facility questionnaire. For example, field testing a health facility questionnaire should involve visits to public health posts, public clinics, private doctors' offices, public hospitals, and private hospitals in both urban and rural areas. Similarly, field testing a school questionnaire should involve visits to public and private schools, primary and secondary schools, and schools in urban and rural areas. Since field testing a facility questionnaire is a major undertaking in its own right, it is probably best to conduct such a field test separately from the field tests of the other questionnaires.

### Rules for Formatting Survey Questionnaires

The formatting of survey questionnaires is not a separate step in the overall survey design process. Rather, it influences how the third, fourth, and fifth steps are carried out. Good questionnaire formatting can make a tremendous difference in the quality of the data collected. This section discusses formatting in detail, making very specific recommendations about how questionnaires should be formatted.<sup>9</sup>

There is, of course, more than one way to format household survey questionnaires. Most of the benefits of good formatting come from selecting a formatting convention and following that convention consistently, rather than choosing the "best" convention from among several possible options. For example, in LSMS questionnaires uppercase and lowercase letters are used to distinguish words spoken aloud during the interview from instructions to the interviewer, but this could be done in other ways, such as using different colors or different fonts. Once a convention is selected, it is extremely important to use it consistently throughout the whole questionnaire. The convention

chosen should be the one that is clearest and most likely to minimize the possibility of errors. The draft questionnaires presented in this book follow the formatting conventions explained in this section, which have been used frequently in past LSMS surveys, with successful results.

Questionnaire format is important because a good format minimizes potential interviewer and data entry errors, which improves the accuracy of the data and reduces the time needed to check the data before making them available to data analysts. The objectives underlying a given survey can occasionally have implications for formatting, so some aspects of formatting will vary from country to country. Even so, almost all of what has been learned about questionnaire format in previous LSMS surveys will be applicable to new surveys. Thus the formatting guidelines presented in this section are recommended for all LSMS and similar multitopic surveys, and for other surveys as well.

### Identifiers

Every person or object for which data are collected in a survey must be uniquely identified. This usually requires two or three separate codes. The first code identifies the household. The second code identifies the person or object of interest, such as an individual household member, a household business, or a plot of land. Sometimes there is a third code, which applies, for example, to all children ever born to each woman in the household or to the assets of each business operated by the household.

Whenever possible, the identification codes for the second or third levels of observation should be preprinted on the questionnaire pages to which they pertain. For example, the individual identification code for each household member should be printed on all pages that collect data on individual household members. This ensures that the codes cannot be omitted and avoids any errors that would occur if the interviewer were to write down the wrong codes. An example of these codes appears in the far left column of Figure 3.2, which presents the short version of the education module.

The importance of adequate identifiers is so obvious that it is hard to believe mistakes can be made, but they can. In one health survey the questionnaire consisted of two sheets of paper stapled together. One contained information on the household, while the other contained information on individuals. In order

FIGURE 3.2: ILLUSTRATION OF INDIVIDUAL IDENTIFICATION AND SKIP CODES (EDUCATION MODULE SHORT VERSION)

I D C O D E	1.	2.	3.	4.	5.	6.	7.	8.	9.							10.	11.
	Have you ever attended school?	Are you currently enrolled in school?	What is the highest grade you have completed in school?	What is the highest diploma you have attained?	Were you enrolled in school during the past 12 months?	In what grade are you currently enrolled in school?	What is the highest diploma you have attained so far?	Is the school you are currently enrolled in public or private?	How much has your household spent on your education in the last 12 months for:							Have you ever repeated a grade of school?	How many times have you repeated a grade of school?
	YES . . 1 NO . . . 2 (»NEXT PERSON)	YES . . 1 (»6) NO . . . 2	PUT CODES FOR DIFFER- ENT GRADES HERE	PUT CODES FOR DIPLOMAS HERE	YES . . 1 (»9) NO . . . 2 (»10)	PUT CODES FOR DIPLOMAS HERE	PUT CODES FOR DIFFER- ENT GRADES HERE	PUBLIC . . 1 PRIVATE SECU- LAR . . . 2 PRIVATE RELI- GIOUS . . 3	A. Tuition and other required fees?	B. Parent Association fees?	C. Uni- forms and other clothing?	D. Text- books?	E. Other educational materials (exercise books, pens, etc.)?	F. Meals, transportation and/or lodging?	G. Other expenses (extra classes, optional fees)?	YES . . 1 NO . . . 2 (»NEXT PERSON)	»»NEXT PERSON  NUMBER OF REPEATED GRADES
1																	
2																	
3																	
4																	
5																	
6																	
7																	
8																	
9																	
10																	
11																	
12																	

to facilitate data entry, the two pages of the questionnaire were separated. Unfortunately, the household identifier was not put on the page for individuals, making it impossible to link the two parts of the survey with each other after the data were entered.

### *Questionnaire Layout*

The LSMS questionnaires are designed so that only one copy of the questionnaire is needed for each household. In contrast, some surveys use one household questionnaire and a separate set of individual questionnaires. This requires that household identification codes be copied perfectly onto all of the individual questionnaires. While perfection is always sought, it is rarely achieved, and separate questionnaires create the risk of improper matching. This is illustrated in the case of the Russian Longitudinal Monitoring Survey. Although care was taken to ensure accurate coding and matching, many errors were introduced. For the first round of the survey, which was held in the summer of 1992, there were 3 percent fewer individual questionnaires than had been expected given the number of household members identified in the household questionnaires. By the summer of 1993, in the third round of the survey, this discrepancy had grown to about 9.5 percent.

Putting all of the information into a single household questionnaire implies the need for a grid of some kind whenever there are two or more of a particular unit of analysis in a household. For example, a household often includes several people, may have several plots of land, and may grow several different crops. The grid typically used in LSMS surveys has questions arranged across the top and units of observation (people, plots, or crops) down the side; in other words, each question is a column and each unit of observation is a row. An example of this is shown in Figure 3.2; note that the identification codes for the units of observation (household members) are printed on the left side of the grid page.

Sometimes the interviewer must fill in the code in the first column, as in Question 2 of Figure 3.8 (which is discussed below), but this practice should be minimized to reduce the possibility of introducing errors when writing down such codes. In the grids for individuals, the lines can be differentiated by alternating shaded and unshaded blocks (as in the draft modules in Volume 3 of this book) or by using a different color for each row or block of rows. This helps an interviewer record the information on the correct line.

Exceptionally large households sometimes have so many members that there are not enough lines in the grids for all household members. In these cases a second copy of the household questionnaire will be required, and care must be taken to ensure that the right household and individual numbers are used. As explained in Chapter 4, a coding scheme is needed to distinguish between the first and second copies of the questionnaire filled out for large households. For example, the individual numbers in the second copy should be changed to start with 13 instead of 1 (assuming that the first questionnaire has room for 12 household members). This is a reasonable approach for large households, but it also introduces a potential source of error; survey designers should set the format of the grids to accommodate as many individuals as is practical. Previous LSMS questionnaires have typically had space for 12–15 individuals.

In cases where the unit of analysis is such that there is only one observation per household (for example, one dwelling per household), the questions pertaining to that unit can be arranged in a single column down the page. One problem with a single column of questions is that much of the page is left blank. To save paper, two or more columns may be put on one page, as long as it is clear that there is no horizontal relationship among the questions in the different columns. An example of this format is provided in Figure 3.3, which shows the first page of Part C of the standard housing module.

### *Fold-Out Roster Page*

The household roster page of the household questionnaire is printed so that it extends to the left of the pages that pertain to individuals in the household. Most importantly, the names of each individual member of the household on the roster page are visible when filling out the other individual-specific pages of the household questionnaire. This has been done four different ways in LSMS surveys, as illustrated by Figure 3.4.

In the method shown in Format 1, the sheets in front of the roster are shorter than the cover, the roster, and the sheets that follow the roster. The most common method is shown in Format 2. The roster sheet is folded out to extend beyond the body of the questionnaire and its covers. In Formats 1 and 2 the roster page is placed behind all of the pages that pertain to individuals, so that the names on the household

FIGURE 3.3: ILLUSTRATION OF PRECODING (PART C OF HOUSING MODULE)

1. Is this dwelling owned by a member of your household?

YES .....1  
NO .....2 (>11)

2. How did your household obtain this dwelling?

PRIVATIZED .....1  
PURCHASED FROM A PRIVATE PERSON .....2  
NEWLY BUILT .....3  
COOPERATIVE ARRANGEMENT .....4  
SWAPPED .....5 (>6)  
INHERITED .....6 (>6)  
OTHER .....7 (>6)

3. How much did you pay for the unit ?

4. If you make installment payments for your dwelling, what is the amount of the installment?

WRITE ZERO IF THE HOUSEHOLD DOES NOT MAKE  
INSTALLMENT PAYMENTS  
AMOUNT (UNITS OF CURRENCY)  
TIME UNIT

5. In what year do you expect to make your last installment payment?

YEAR

6. Do you have legal title to the land or any document that shows ownership?

YES .....1  
NO .....2

TIME UNITS:	DAY.....3	MONTH.....6	YEAR..9
	WEEK.....4	QUARTER.....7	
	FORTNIGHT..5	HALF YEAR...8	

7. Do you have legal title to the dwelling or any document that shows ownership?

YES .....1  
NO .....2

8. What type of title is it?

FULL LEGAL TITLE, REGISTERED ..1  
LEGAL TITLE, UNREGISTERED .....2  
PURCHASE RECEIPT .....3  
OTHER .....4

9. Which person holds the title or document to this dwelling?

WRITE ID CODE OF THIS PERSON FROM THE ROSTER

1ST ID CODE:   
2ND ID CODE:

10. Could you sell this dwelling if you wanted to?

YES .....1  
NO .....2 (>13)

11. If you sold this dwelling today how much would you receive for it?

AMOUNT (UNITS OF CURRENCY)

12. Estimate, please, the amount of money you could receive as rent if you let this dwelling to another person?

AMOUNT (UNITS OF CURRENCY)   
TIME UNIT

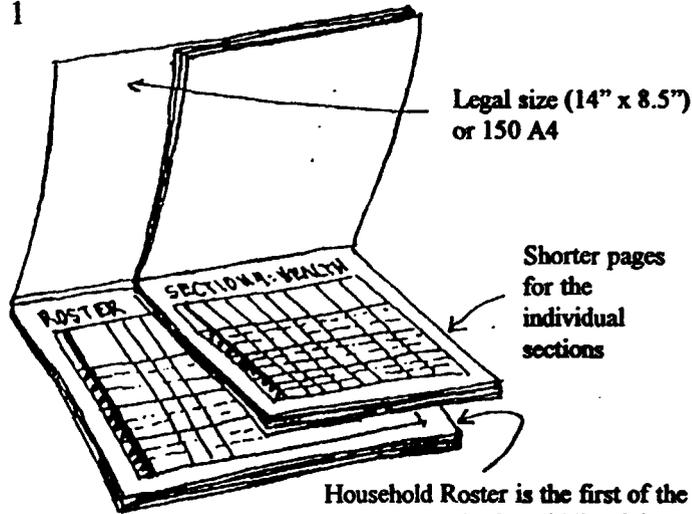
»» QUESTION 28

13. Do you rent this dwelling for goods, services or cash?

YES .....1  
NO .....2 (>26)

FIGURE 3.4: ROSTER ARRANGEMENTS

Format 1

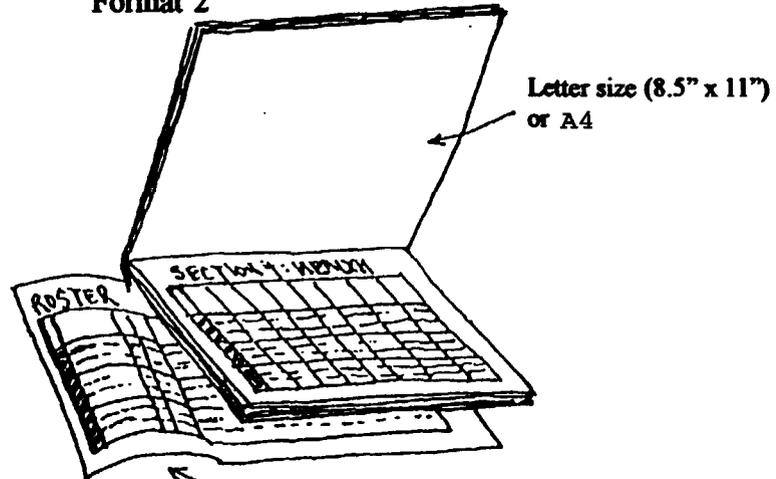


Legal size (14" x 8.5")  
or 150 A4

Shorter pages  
for the  
individual  
sections

Household Roster is the first of the  
longer pages in the middle of the  
questionnaire

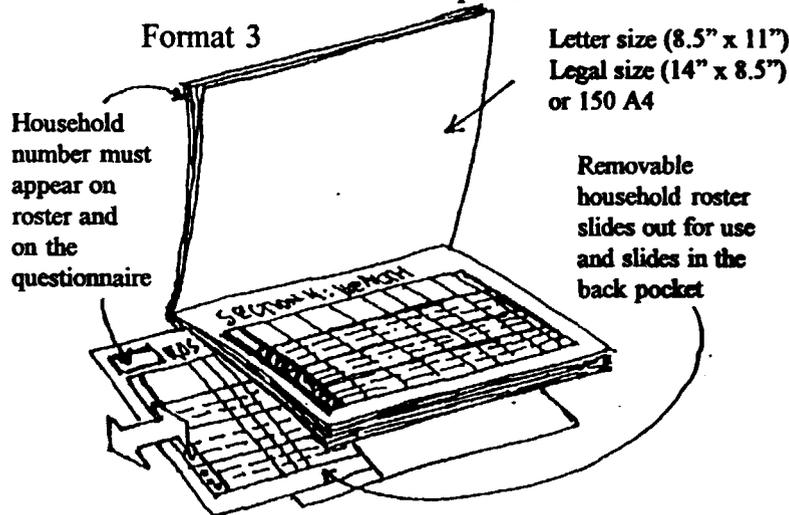
Format 2



Letter size (8.5" x 11")  
or A4

Household Roster on a wider  
sheet; folds out from the back page

Format 3

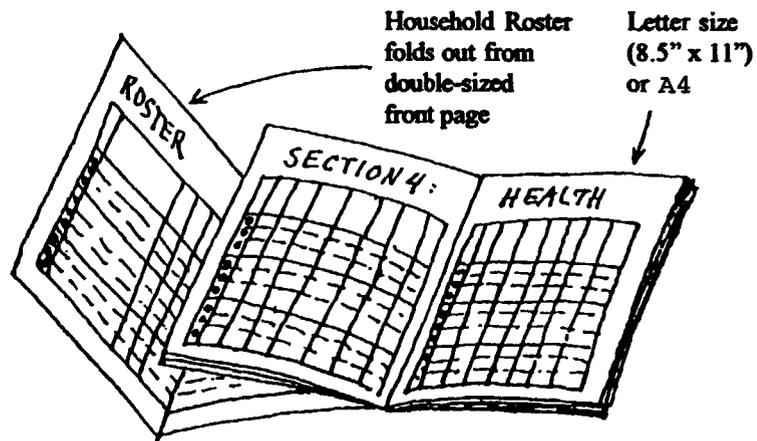


Letter size (8.5" x 11")  
Legal size (14" x 8.5")  
or 150 A4

Removable  
household roster  
slides out for use  
and slides in the  
back pocket

Household  
number must  
appear on  
roster and  
on the  
questionnaire

Format 4



Household Roster  
folds out from  
double-sized  
front page

Letter size  
(8.5" x 11")  
or A4

In all formats, choose binding to make questionnaire open flat. ID codes appear on the roster and on each individual page. Lines on the roster must be aligned with the pages in the questionnaire.

roster page are visible whenever individual questions are asked.

An innovation in the Kagera Health and Development Survey in Tanzania was to make the roster page a removable card, as shown in Format 3. This was useful because the survey was designed to be administered four times—every six months for two years—to the same households. The roster card was inserted into a pocket in the back of the questionnaire in the first round of the survey. When the second round started, the roster card was removed from the first questionnaire and placed in the back pocket of the second questionnaire. In this way, individuals retained the same identification codes in each round. A few follow-up questions guaranteed that individuals who moved in or out of the household or were born or died between rounds were counted appropriately. In four rounds of interviews conducted over two years for 800 households, none of the roster cards was lost. However, this success may reflect the intensive supervision carried out by the organizers of that survey, as well as the relatively small sample size. This option should probably not be used in situations with significant quality control problems.

Format 4 was used in the Tunisia questionnaire. In this format each page is oriented as “portrait” (a vertical page) rather than as “landscape” (a horizontal page) and is spiral-bound so that it opens flat. Each questionnaire page then consists of the full 11 x 17 inches of the two-page spread. The roster folds out to the left. In all four cases the line for each individual member of the household on the roster page is aligned with the corresponding lines on the other individual-specific pages of the household questionnaire.

A final point regarding the fold-out roster page is that it may be useful to have more than one such page per questionnaire. A fold-out roster will be useful whenever there are several pages of questions for the same level of analysis and especially when there are many rows on the grid. For example, in the agricultural module one might make rosters for crops grown or for plots of land. A fold-out roster page would be particularly helpful for the household enterprise module.

#### **Precoding**

All of the potential responses to almost all of the questions in the questionnaire should be given code numbers so that the interviewer records only code numbers, as opposed to words or phrases, on the

questionnaire. In most cases these response codes should be printed directly in the box where the question appears, or next to the question if there is no box around it. Where the list of codes is lengthy and applies to several questions, it should be placed in a special box on the border of each page for which it is needed. Alternatively, if a list is very long it can be printed on the back of the preceding page (making it visible when the interviewer fills out the page in question). An example of a box on a border of a page is the time unit box shown at the bottom of Figure 3.3.

In past LSMS surveys fewer than a dozen questions on the household questionnaire have required the interviewer to write down words or phrases that are given codes, usually by someone else, after the interview. Precoding allows the data to be entered into the computer straight from the completed questionnaire, thus eliminating the time-consuming and error-prone step of transcribing codes onto data entry sheets.

Precoding requires that response codes be clear, simple, and mutually exclusive, that they exhaust all likely answers, that respondents will not all provide the same response, and that none of the codes apply to only a handful of respondents. Designing adequate response codes requires extensive knowledge of the phenomenon being studied as well as careful field testing. A standard technique to ensure that the codes are mutually exclusive is to add a qualifier where more than one answer could apply—asking, for example, “What was the *main* reason for dropping out of school?” Other standard qualifiers are “What was the first (or last, or principal) reason for...?” Alternatively, spaces can be provided for multiple responses, with an instruction to code all responses (up to, say, the two or three most important) that apply.

A standard technique to ensure that codes encompass all possible answers is to add an “other (specify \_\_\_\_\_)” code to questions for which an explicit enumeration of each possible response is impossible or inconvenient. In past LSMS surveys the detailed answers were almost never coded, so analysts usually put all “other” responses into a single residual category. One way to increase the probability that the information recorded in the “other (specify \_\_\_\_\_)” answers will be used at a later date is to enter it (as text) into the computer, without assigning any codes to the responses. This allows analysts to code any answers that were not precoded in the data released to

the public. It also allows the designers of subsequent surveys in the same country to review the answers that were written in (especially in cases in which a significant percentage of the responses were coded “other”) and to modify their coding lists accordingly. In particular, if most of the “other” responses fall into a single, well-defined category, this category should have its own code in any subsequent survey.

There is, of course, a limit to the kind of material that can be covered even by well-designed, precoded questions. But this limit may be less of a disadvantage than it first appears. Because most analyses of LSMS surveys use sophisticated quantitative techniques, it is difficult for these analysts to make use of the exploratory, qualitative information gathered in open-ended questions. So even if such questions were asked, the answers to these questions would not be used much in analysis. If it is clear that some analysts do need extensive information of an exploratory, qualitative nature, the designers of a prospective survey may wish to adopt a different data collection instrument or even a new research technique. See Chapter 25 for a thorough discussion of qualitative data collection alternatives.

#### *Verbatim Questions with Simple Answers*

All questions in LSMS surveys are written out in their entirety and are meant to be read out verbatim by the interviewer. This is done to ensure that questions are asked in a uniform way, since different wordings may elicit different responses. For example, the answers that a respondent gives to “Can you read?” and to “Can you read a newspaper or magazine?” will probably be somewhat different. Other changes may subtly alter the time period referred to, as in the change from “Have you worked since you were married?” to “Did you work after you were married?” Scott and others (1988) discuss some rigorous field experiments that compared such verbatim questionnaires with questionnaires in which the topic was given for each question but the exact wording was not. When the questionnaire that did not contain the exact wording was used, 7 to 20 times more errors occurred than when the verbatim questionnaire was used.

When choosing the wording of questions, it is important to use terms that reflect the language as it is commonly spoken. Using language that is too formal or academic will make the interview stilted and

unnatural. For example, “Did you spend any time doing housework?” followed, if necessary, by “...such as cooking, mending, doing laundry, or cleaning?” is better than “Did you spend any time engaged in domestic labor, for example, preparing food, repairing clothes, cleaning clothes, or cleaning house?” It is not always easy to find terms that are simple, short, and yet precise, but that should always be the goal.

In most cases the interviewer reads the question aloud and marks the questionnaire with the code for the answer given by the respondent. For example, for the question, “Are you currently enrolled in school?” the interviewer writes down a 1 for “yes” or a 2 for “no.” For some questions the response categories are part of the question—for example, “Is the school you are currently enrolled in public or private?” There may also be a few questions for which the wording of respondents’ answers may vary even though the meaning is the same. The best thing to do in such cases is to have the interviewer read out all of the response categories. For example, in Question 4 of Figure 3.5, after reading “Compared to your health one year ago, would you say that your health is...” the interviewer should read the responses “much better now,” “somewhat better now,” “about the same,” “somewhat worse,” and “much worse.” If necessary, the interviewer can explain the differences between the various response categories. However, the reading out of response categories should be used as little as possible, because respondents may not listen to the full list before answering, which can lead to errors.

The answers to the questions must be kept simple. This means that additional filter questions are often needed. Adding enough filter questions to ensure simple answers can make the number of questions and skips seem high. Many survey designers are tempted to shorten the questionnaire or simplify the skip pattern in a way that results in complex questions and answers. This should be avoided since it will confuse some respondents and is unlikely to save time.

Survey designers yielded to this temptation in the agricultural module of the 1987–88 Ghana LSMS survey. In that module the following question was asked: “Do you or the members of your household have the right to sell all or part of their land to someone else if they wish?” The precoded answers (which were not read out to the respondents) were “Yes,” “No,” “Only after consulting family members who are not household members,” and “Only after consulting the chief or



the village elders.” It is not clear whether the respondents could distinguish between the simple yes answer and the yes answer qualified by the need for consultation. Thus a different formulation might have been better. The question could have been left as is but using only simple “yes” and “no” codes. Then the interviewer could have put a second question to those who answered “yes,” worded as follows: “Do you need to consult with anyone outside the household before selling the land?” The response codes would be “Yes” and “No.” Then a third question would be put to those who answered “yes” to the second: “Whom must you consult?” The response codes for this question would be for “family member,” “village elders,” and other appropriate categories. This formulation would have made the questionnaire longer in terms of the number of questions but would probably not have increased the interview time since some sort of probing probably occurred in the Ghana LSMS when the “yes” answer was given. More importantly, keeping questions and answers simple makes the interpretation of the data much clearer.

#### *Skip Codes*

Skip codes are used extensively in LSMS questionnaires. Skip codes tell the interviewer which question to proceed to after finishing the current question. Some skip codes apply only when a particular answer is given. In such cases an arrow and the number of the question to skip to are positioned in parentheses next to or below the individual response to which the code applies. An example is given in Question 2 of Figure 3.2. If the answer to Question 2 is “yes,” the interviewer should skip Questions 3, 4, and 5 and proceed to Question 6. If the answer to Question 2 is “no,” the interviewer should proceed to Question 3. In Question 1 a similar construction is used, but when the answer is “no” the interviewer is instructed to skip all the remaining questions in the module for this respondent and proceed to interview the next person.

Another kind of skip instruction applies regardless of the response given to the question. When an arrow and a question number or instruction are placed in a box separate from the response codes, the skip instruction contained in the box applies regardless of what answer is given. An example of this is given in Question 10 of Figure 3.5.

There are several advantages to extensive, explicit skip codes. Interviewers do not have to make decisions themselves, nor do they need to remember complicat-

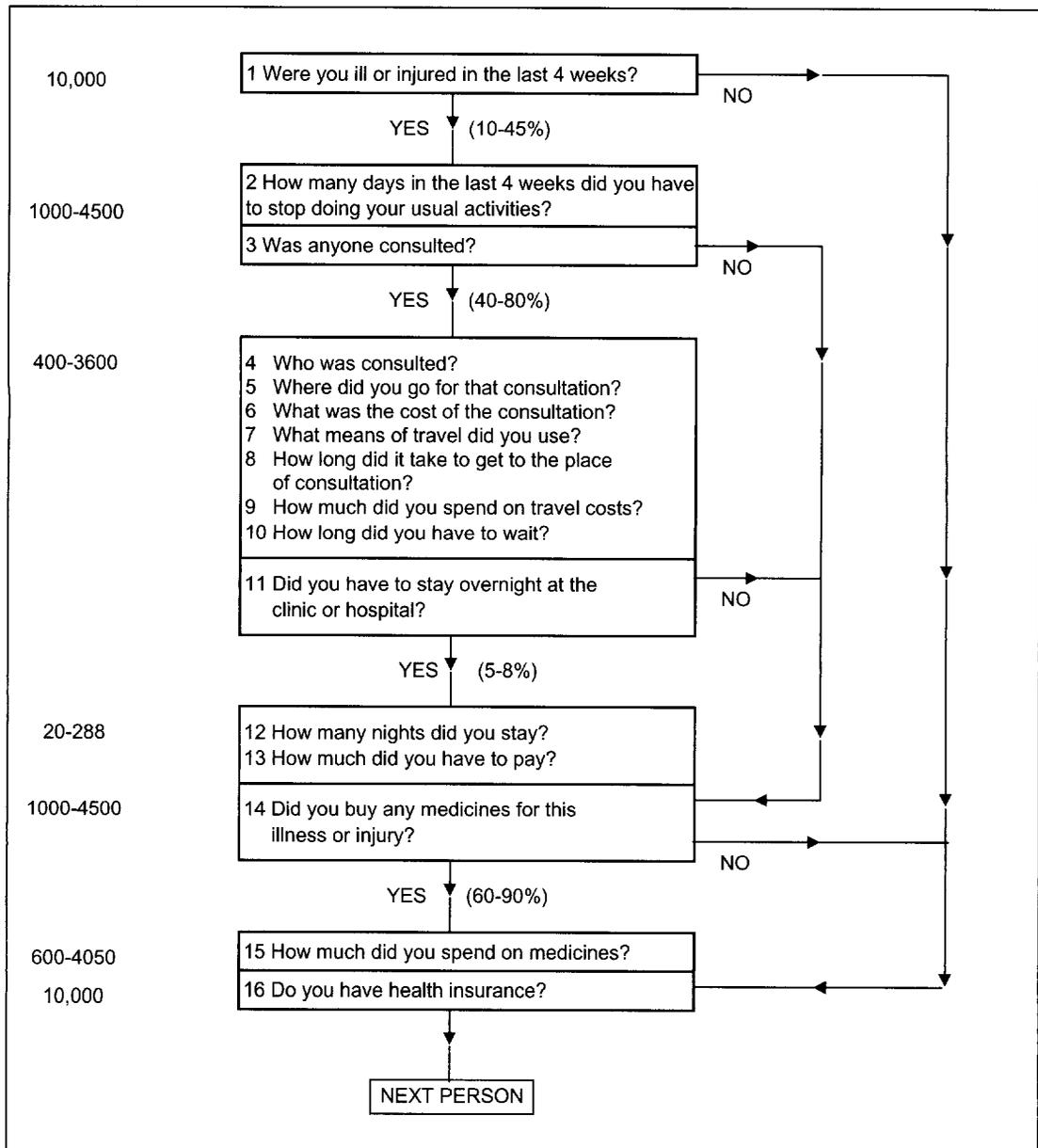
ed rules printed in the manual rather than on the questionnaire. This helps ensure that instructions will be followed uniformly. Well-placed skip codes ensure that inapplicable questions are not asked. (Asking inapplicable questions irritates respondents, wastes interview time, and confuses data analysis.) Finally, explicit skip codes imply that a “not applicable” code is almost never used in LSMS questionnaires.

One way to check skip codes is to develop a flow chart of the questions in each module. Flow charts are useful both for checking the logic of the questionnaire and for training interviewers. Figure 3.6 presents a flow chart of a typical health module used in past LSMS surveys (which differs in several notable ways from the health module presented in Volume 3). The proportions of people who answer yes at each branch are recorded based on results from several previous LSMS surveys. The numbers of individuals that would be asked each set of questions are shown on the left, assuming a base of 10,000 individuals in the sample. The flow chart makes it easy to check whether the skip patterns lead people through the module correctly. For example, it is possible to check that the question on health insurance is asked of all household members, not just of those who are ill. Analyzing the whole household in this way gives survey designers a better sense of the likely length of time it will take to complete each interview than does the number of pages or number of questions in the questionnaire, because many questions will be skipped for many individuals. (For further discussion of the length of the questionnaire see the second section of this chapter.)

#### *Case Conventions*

Everything that the interviewer should read aloud should be written in lowercase letters. Instructions to the interviewer should always be written in uppercase letters.<sup>10</sup> Answer codes should also be written in uppercase, unless they are to be read aloud to the respondent. This makes it easy to include instructions on the questionnaire as opposed to relying on the interviewers’ memory of the manual or of instructions that they were given during their training. In Figure 3.5 instructions to the interviewer are printed in Questions 1, 2, 4, and 5. These are in uppercase, as are the answer codes in Questions 1 and 5. (The answer codes in Question 4 are in lowercase because they are to be read aloud to the respondent.)

FIGURE 3.6: FLOW CHART OF HEALTH MODULE USED IN PREVIOUS LSMS SURVEYS



*Enumeration of Lists*

There are two methods of gathering information about long lists of items. A typical LSMS questionnaire may use either method depending on particular circumstances.

Consider the case in which one expects that a large proportion of the items on the list will apply to most households. For each item on this list a line is put in the grid and the name and code number of the item is printed on the questionnaire. This approach is

used in the consumption module, as shown in Figure 3.7. Although several dozen items are included, it is expected that most households will have consumed many of them. The first question is “Has your household consumed [FOOD] during the past 12 months?” The interviewer first goes down the whole list asking this “yes or no” question. Then the interviewer returns to the first item that was consumed and asks all the follow-up questions for that item before proceeding to the next item. The complete

**FIGURE 3.7: ILLUSTRATION OF CLOSE- ENDED LIST (PART B OF CONSUMPTION MODULE)**

1.	PURCHASES SINCE LAST VISIT			PURCHASES TYPICAL MONTH			HOME PRODUCTION			GIFTS
	2.	3.	4.	5.	6.	7.	8.	9.	10.	
<p>In the following questions, I want to ask about all purchases made for your household, regardless of which person made them.</p> <p>Has your household consumed [FOOD] during the past 12 months? Please exclude from your answer any [ITEM] purchased for processing or resale in a household enterprise.</p> <p>PUT A CHECK (✓) IN THE APPROPRIATE BOX FOR EACH FOOD ITEM. IF THE ANSWER TO Q.1 IS YES, ASK Q.2-13.</p>	<p>Have the members of your household bought any [FOOD] since my last visit, that is since [DAY/DATE]?</p> <p>YES . 1 NO . . 2 (»5)</p>	<p>How much did you pay in total?</p> <p>CURRENCY</p>	<p>How much did you buy?</p> <p>AMT UNIT</p>	<p>How many months in the past 12 months did your household purchase [FOOD]?</p> <p>IF NONE WRITE ZERO, » 7</p> <p>MONTHS</p>	<p>How much do you usually spend on [FOOD] in one of the months that you purchase [FOOD]?</p> <p>CURRENCY</p>	<p>How many months in the past 12 months did your household consume [FOOD] that you grew or produced at home?</p> <p>IF NONE WRITE ZERO, » 10</p> <p>MONTHS</p>	<p>How much did you consume in a typical month?</p> <p>AMT UNIT</p>	<p>What was the value of the [FOOD] you consumed in a typical month from your own production?</p> <p>CURRENCY</p>	<p>What is the total value of the [FOOD] consumed that you received as a gift over the past 12 months?</p> <p>IF NONE, WRITE ZERO</p> <p>CURRENCY</p>	
<p>NO YES CODE</p>										
Wheat (grain)										
Wheat (flour or maida)										
Maize (flour or grain)										
Jawar/Bajra										
Fine rice (basmati)										
Coarse rice										
Other grains/cereals										
Gram										
Dal										
Groundnuts										
Liquid vegetable oils (dalda)										
Ghee, Desi ghee										
Fresh milk										

**UNIT CODES:  
USE CODES  
WITH STAR  
WHENEVER  
POSSIBLE**

KILO\*....1  
GRAM\*....2  
POUND\*...3  
OUNCE\*...4  
LITER\*...5  
CUP\*.....6  
PINT\*.....7  
QUART\*...8  
GALLON\*..9  
BUNCH...10  
PECK....11  
BUSHEL..12  
TIN.....13  
PIECES..14  
DOZENS..15  
BOTTLES.16

FIGURE 3.7: ILLUSTRATION OF CLOSE- ENDED LIST (PART B OF CONSUMPTION MODULE)

1.	PURCHASES SINCE LAST VISIT			PURCHASES TYPICAL MONTH			HOME PRODUCTION			GIFTS			
	NO	YES	CODE	CURRENCY	AMT	UNIT	MONTHS	CURRENCY	MONTHS	AMT	UNIT	CURRENCY	CURRENCY
<p>In the following questions, I want to ask about all purchases made for your household, regardless of which person made them.</p> <p>Has your household consumed [FOOD] during the past 12 months? Please exclude from your answer any [ITEM] purchased for processing or resale in a household enterprise.</p> <p>PUT A CHECK (✓) IN THE APPROPRIATE BOX FOR EACH FOOD ITEM. IF THE ANSWER TO Q.1 IS YES, ASK Q.2-13.</p>	<p>2. Have the members of your household bought any [FOOD] since my last visit, that is since [DAY/DATE]?</p> <p>YES . 1 NO . . 2 (&gt;5)</p>			<p>3. How much did you pay in total?</p> <p>CURRENCY</p>		<p>4. How much did you buy?</p> <p>AMT UNIT</p>	<p>5. How many months in the past 12 months did your household purchase [FOOD]?</p> <p>MONTHS</p> <p>IF NONE WRITE ZERO, » 7</p>	<p>6. How much do you usually spend on [FOOD] in one of the months that you purchase [FOOD]?</p> <p>CURRENCY</p>	<p>7. How many months in the past 12 months did your household consume [FOOD] that you grew or produced at home?</p> <p>MONTHS</p> <p>IF NONE WRITE ZERO, » 10</p>	<p>8. How much did you consume in a typical month?</p> <p>AMT UNIT</p>		<p>9. What was the value of the [FOOD] you consumed in a typical month from your own production?</p> <p>CURRENCY</p>	<p>10. What is the total value of the [FOOD] consumed that you received as a gift over the past 12 months?</p> <p>CURRENCY</p> <p>IF NONE, WRITE ZERO</p>
Yogurt and Lassi			14										
Milk Powder			15										
Baby Formula			16										
Sugar (refined)			17										

- UNIT CODES:  
USE CODES  
WITH STAR  
WHENEVER  
POSSIBLE**
- KILO\* . . . . 1
  - GRAM\* . . . . 2
  - POUND\* . . . . 3
  - OUNCE\* . . . . 4
  - LITER\* . . . . 5
  - CUP\* . . . . 6
  - PINT\* . . . . 7
  - QUART\* . . . . 8
  - GALLON\* . . . . 9
  - BUNCH . . . . 10
  - PECK . . . . 11
  - BUSHEL . . . . 12
  - TIN . . . . 13
  - PIECES . . . . 14
  - DOZENS . . . . 15
  - BOTTLES . . . . 16

enumeration of items consumed is done before asking the follow-up questions so that respondents will not be tempted to say that they have not consumed something in order to shorten the interview by avoiding the follow-up questions. This temptation is prevented because the enumeration is done before the respondent finds out that there will be follow-up questions on each item enumerated.

A second approach is useful when it is expected that only a few of many possible items will pertain to any one household. Consider Figure 3.8. The large grid on the right contains lines for several durable goods owned by the household, but these are not pre-coded. Rather, the respondent is asked, using the small grid on the left, whether the household owns certain durable goods. In this example 12 durable goods are considered, but in some cases 20–30 goods have been listed. Most households own only a few durable goods. For all durable goods owned by the household, the interviewer lists the name and the code number in the large grid to the right in Figure 3.8, and asks a series of questions about each good. If the household owns two or more of the same durable good, one line is filled out for each good owned.

#### *Probe Questions*

There are some kinds of information that respondents may accidentally not provide. In such cases the questionnaire includes instructions to the interviewer to ask further “probing” questions on the subject. An example of this is Question 9 of Figure 3.1. Suggested probing questions are usually included in the interviewers’ manual and occasionally included in the questionnaire itself. Probe questions are often used to ensure that all items in a respondent-determined list have been reported to the interviewer, or to ensure that the respondent’s answer is properly classified by the interviewer. Interviewers are also asked to probe for answers to questions that ask “how much...?” (This kind of question is commonly found in the consumption, agriculture, and household enterprise modules.) Interviewers should be thoroughly trained to ensure that they fully understand what information to probe for, and how to do so.

Because the interviewer is trained and instructed to probe for information, there should be very few answers of “don’t know” and thus very few codes for “don’t know” in the questionnaire. In the exceptional case when even a sound interviewing technique does

not produce an answer, the interviewer is instructed (in the interviewers’ manual and in training) to write “DK” (for “don’t know”) in the space reserved for an answer code. Such responses are given a special non-numeric code in the data entry program. The end result for analysis is much the same as having a “don’t know” code for each question. However, this system has the advantage that it discourages interviewers from accepting “don’t know” answers too easily, which they may be tempted to do to speed up the interview. Moreover, the special non-numeric code for such responses is glaringly obvious when the supervisor reviews the questionnaire.

#### *Letting Respondents Choose Units*

For many questions that involve payments or quantities, respondents are allowed to give their answers in whatever units they find most convenient. Examples of this are found in Figure 3.3. In Questions 4 and 12 the code of the time unit in which the respondent replies is placed in the box marked “time unit.” The codes are provided in a box at the bottom of the page.

Allowing the respondent to select the time unit means that transactions are expressed in the units in which they normally occur, which may differ from household to household or from person to person. This avoids inaccuracies in conversion. For example, a person paid \$510 per week will respond precisely if allowed to respond on a per-week basis. If forced to respond in terms of dollars per month, the respondent might round the figure down to \$500 for ease of multiplication and calculate each month as being equivalent to four weeks. The annualized figure would thus become \$24,000 instead of the \$26,520 that would be reported if the respondent were allowed to report on a per-week basis and the data analyst then calculated the respondent’s annual rate from that answer.

Of course, data analysis is always slightly more complicated when respondents’ answers must be converted in order to arrive at annualized figures, but, since a computer can easily do this, this disadvantage is trivial. However, it is very important to ensure that, where necessary, the questionnaire explicitly asks the respondent how many times per year the payments are made. For example, a worker who reports a daily wage rate may be employed only intermittently. In this case, the questionnaire should ask the respondent how many weeks or months he or she has worked during the preceding 12 months (see Chapter 9 for details).

FIGURE 3.8: ILLUSTRATION OF OPEN-ENDED LIST (PART E OF CONSUMPTION MODULE)

1. Does your household own any of the following items?

DETERMINE WHICH DURABLES THE HOUSEHOLD OWNS BY ASKING Q.1. FOR EACH DURABLE OWNED, WRITE THE DESCRIPTION AND CODE IN THE SPACE PROVIDED UNDER Q.2, AND PROCEED TO ASK Q.3-7 FOR EACH ITEM.

ITEM	CODE	YES	NO
Stove	201		
Refrigerator	202		
Washing Machine	203		
Sewing/knitting machine	204		
Fan	205		
Television	206		
Video player	207		
Tape player/CD player	208		
Camera, video camera	209		
Bicycle	210		
Motorcycle/scooter	211		
Car or truck	212		

I T E M	2. LIST ALL THE ITEMS OWNED BY THE HOUSEHOLD, THEN PROCEED TO ASK Q.3-7.	3. How many years ago did you acquire this [ITEM]?	4. Did you purchase it or receive it as a gift or payment for services?	5. How much did you pay for it?	6. How much was it worth when you received it?	7. If you wanted to sell this [ITEM] today, how much would you receive?
	DESCRIPTION	CODE	YEARS	PURCHASE . . 1 GIFT OR PAYMENT . . 2 (>6) CURRENCY	CURRENCY	CURRENCY
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						

**Table 3.1 Units of Quantity Used in Ghana, 1987–88**

Unit	Code
Pound	*1
Kilogram	*2
Ton	*3
Minibag	*4
Maxibag	*5
Sheet	6
Basket	7
Bowl	8
American tin	*9
Tree	10
Stick	11
Bundle	12
Barrel	13
Liter	*14
Gallon	*15
Beer bottle	*16
Bunch	17
Nut	18
Fruit	19
Log	20
Box	21
All	22

Note: It is preferable to use the unit codes marked by (\*) whenever possible.

Source: Ghana LSMS survey (1987–88).

A particular place in the questionnaire where it is useful to allow respondents to choose their own units is in the “quantities produced” questions in the agriculture module. In Ghana, for example, respondents were allowed to give answers in 22 different kinds of units (Table 3.1). A serious problem for analysts who want to convert these different quantities to a single standard unit is that only about half of the units used in this example were standardized, and some of the standardized units were local terms (such as minibag and maxibag) that would be unknown to anyone not familiar with farming in Ghana.<sup>11</sup> In the case of standardized local units, the survey team should ensure that such terms are defined (in terms of international standardized units) in a basic information document that includes all of the information that data users will need to analyze the data.

#### **Respondent Codes**

It is sometimes useful to know who is answering a certain section of the questionnaire. In general, each household member should answer for himself or herself, but this is not always possible. For example, a household member may be away during the entire week when the field team is working in his or her community. To indi-

cate who is providing the information, a question can be inserted that asks the interviewer whether the person is answering for himself or herself. If someone else is providing the information, the interviewer should fill in the identification code of that person. An example of this is shown in Figure 3.5. This information is useful because a proxy respondent may give less accurate information than the individual who is actually involved in the activity in question. For example, one household member may not know the exact salary of another. Therefore, some analysts may wish to identify any possible biases introduced by the proxy respondents or to omit their responses altogether.

#### **Cardstock Covers**

LSMS questionnaires are usually printed with cardstock covers—covers made of very thin cardboard similar to the cardboard used in file folders. In some past surveys it was decided not to use these covers because of their added cost, but this led to the problem that the front and back pages of the questionnaire occasionally came loose. Since the front page usually carries the key household identifier information and the back page sometimes contains the household roster, any such loss is likely to render the rest of the questionnaire useless. Thus cardstock covers are well worth their cost.

#### **Identifying Sections**

The household questionnaire contained in a prototypical full LSMS survey can be very bulky. The Nepal questionnaire, for example, had 70 pages. Therefore, it is useful to devise some ways to make it easy for readers to find their way around in these questionnaires. A few ideas are listed here, and there may well be more. First, it is useful to have page numbers on each page and a table of contents listing the sections (and their page numbers) at the beginning or end of the household questionnaire. Second, some inexpensive graphic techniques can be used to divide the questionnaire into smaller parts. For example, some sections of the questionnaire can be printed on different colored paper or in different colored inks, or sheets of colored paper can be inserted between major portions of the questionnaire. It is also possible to print short, dark bars at the edge of each page, with the placement of these bars on the page being the same within each module but lower down (if on the vertical edge) or further to the right (if on the bottom edge) in each

successive module. Using just one or a few of these techniques will be sufficient. The questionnaire should not become too colorful or complicated.

### *Legibility and Spacing*

There is an art to laying out the grids for a questionnaire. The lettering must be large enough to read, which is sometimes difficult to accomplish in the compact structure of the grid. Legibility is especially important, as interviews often take place under poor lighting conditions, such as outdoors at dusk or after dark in homes dimly lit with lanterns, oil lamps, or candles. The good print quality now available from laser printers helps, but poor legibility is an ongoing complaint among interviewers.

There must also be enough white (empty) space in the layout of the questionnaire. Whenever the answer will be coded later, a generous space should be allowed to write out fully the information required, such as the person's name, the name of the school attended by the respondent, and the respondent's occupation. In other places, judicious use of white space makes the questionnaire easier to read or less confusing than a questionnaire in which every page is crowded with print.

In fact, in this book, the fonts used in Volume 3 are probably too small. This is necessary for Volume 3 to show how typical questionnaire pages should appear. In an actual questionnaire, the size of the pages usually will be somewhat larger than the pages in this book, and the font size should be increased by a similar proportion.

### *Software for the Questionnaire Layout*

Many of the most common word processing and graphics software packages are adequate for producing questionnaire page layouts, and LSMS questionnaires have been produced using several different software packages. The modules in Volume 3 (the electronic versions of which are available to readers in the CD-ROM enclosed in the volume) were produced in Microsoft Excel, for two reasons. First, Excel is widely available. Second, spreadsheet software is better than word processing software at dealing with the long horizontal format of groups of questions on a single topic that are spread across several pages. Regardless of the software used, it is now much simpler and cheaper to make revisions between the various drafts of the modules than it was in the days when graphic artists had to draw each page by hand. The computerized approach

also simplifies translations, as the verbal parts can be overwritten in the local language, leaving intact the skip codes, response codes, and general format.

## **Appendix 3.1 Common Gaps and Overlaps**

This appendix provides a list of the modules that should be checked for gaps and overlaps with respect to the information that they collect. This list is not meant to be exhaustive because household questionnaires of different configurations will be subject to different risks of gaps and overlaps and because there are so many possibilities that it is difficult to list them all. However, some of the most common and important issues are mentioned here. Many more are mentioned in the relevant chapters of this book.

### *Consumption*

Consumption information usually comes from several different modules of the household questionnaire. See the discussion in Chapter 5 on the different components of consumption and the modules in which those components are typically collected.

### *Income*

Information on household income is gathered in the following modules: employment, household enterprise, agriculture, and transfers and other nonlabor income. It is sometimes also collected in the housing and savings modules. It is important to review the questionnaire as a whole to make sure that it accounts for all possible sources of income. In particular, questions about income from any rental property could be placed in the transfers and other nonlabor income module, on the assets page of the savings module, or, if the income comes from renting out a portion of the household's primary dwelling, in the housing module.

### *Wealth*

Information on household assets is collected in several modules. The housing module gathers information on the household's principal residence. The household enterprise module gathers information on equipment and land associated with each household enterprise, and on the stocks of inputs and outputs used in each enterprise. The agricultural module gathers information on land, equipment, and livestock. The savings module collects information on other properties and financial assets, and the durable goods submodule of

the consumption module collects data on the household's durable goods. Finally, the credit module gathers information on the household's liabilities.

#### **Credit**

Credit information is collected in several modules, including the modules for housing, consumption, savings, agriculture, and household businesses. There is also a separate credit module. Chapter 21 introduces the credit module and clarifies gaps and overlaps in credit.

#### **Mortgages**

Information on any mortgages that a household might hold can be gathered either in the credit module or in the housing, agriculture, and household enterprise modules.

#### **Employment**

Analysts often need to know how many hours each household member works in the household's enterprises and in its agricultural activities as well as hours worked in employment outside the household. In previous LSMS surveys, all of this information was collected in the employment module. As explained in Chapter 9 (and Chapters 18 and 19), this book recommends collecting data on household members' days and hours of work in household enterprises and agricultural activities in the household enterprise and agriculture modules, respectively, while continuing to ask about the number of hours worked in wage employment in the employment module. However, some survey designers may decide not to include the household enterprise and agriculture modules. In such cases information on the number of hours spent working on these activities must be collected in the employment module.

#### **Vaccination**

If the survey includes a fertility module, questions about vaccination should usually be placed in the fertility module so that this information can be collected not only for children who currently live in the household but also for children who have died or moved to another household. If there is no fertility module or the fertility module does not include all women of childbearing age, vaccination information on children living in the household can be collected in the anthropology module. Another alternative is to gather this

information in the health module, which includes questions on vaccinations in Part C of the standard health questionnaire.

#### **Domestic Housework**

Some previous LSMS surveys have collected information on how much time household members spend doing housework (such as cooking, cleaning, and childcare) in the employment module, usually asking only one question. If a time use module is included in a questionnaire, there is no reason to ask questions about housework in the employment module. However, because the time use module is very long, it is unlikely to be used in most LSMS-type multitopic surveys. If the time use module is not included but survey designers want to gather a small amount of information on, for example, the number of hours spent on housework during the previous seven days, one or two questions can be added to the employment module. (See Chapter 9 for further discussion of this issue.)

#### **Notes**

The authors would like to express their gratitude to Jere Behrman, Lawrence Haddad, Courtney Harold, John Hoddinott, Alberto Martini, and Raylynn Oliver for comments on an earlier draft.

1. Survey designers occasionally collect redundant information as a cross-check on other data. For example, most previous LSMS surveys have recorded both the age (in years) and the date of birth of each household member. This is done to verify the accuracy of the age variable.
2. This assumes that a two-stage sample is used. In the case of a three-stage sample, the secondary sampling unit is more pertinent. Generally, the penultimate sampling unit is the appropriate unit for collecting community data.
3. Issues concerning the order of the questions within each module are discussed in the topic-specific chapters in Parts 2 and 3 of this book. For a general discussion of ordering questions in household surveys see United Nations (1985) and Frey and Oishi (1995).
4. The short version of the health module presented in Chapter 9 does not ask for particularly sensitive information, but the standard and extended versions ask detailed questions about health status and health behavior (including drinking and smoking) that can be sensitive. If either the standard or the long module is used, health should not be one of the first modules in the questionnaire.
5. The questions in the household enterprise module that refer to "the past 14 days" can be reworded as "since my last visit" if the

second half of the questionnaire is administered two weeks after the interviewer's first visit.

6. For example, the education module asks questions such as "What grade is [..NAME..] enrolled in?" For this question, the range of acceptable values in the data set is precisely defined. Moreover, it is also related to other information such as the degree obtained and the age of the student. (For example, a six-year-old should not be in secondary school.) In the consumption module, however, a wide range of values might be found for a question such as "How much did you spend on rice in the last two weeks?" which implies that fewer consistency checks are possible.

7. This section is a slightly modified version of the discussion on translating and field testing found in Chapter 3 of Grosh and Munoz (1996).

8. An alternative approach is to stretch the reference periods during the field test. For instance, instead of asking "Have you been ill or injured during the past 30 days?" as in the actual survey, it may be expedient to ask "Have you been ill or injured during the past 12 months?" or "When was the last time you were ill or injured?" This approach will simplify the logistics of finding enough people to try out the module but will not test very precisely whether the respondents find it difficult to recall the information, since the recall period used in the field test will be longer than the period used in the final questionnaire.

9. This section is a slightly modified version of the discussion of questionnaire formatting found in Chapter 3 of Grosh and Muñoz (1996).

10. For languages that do not have uppercase and lowercase, another way should be found to distinguish instructions from questions. It may be possible to use italics, bold, a different font, or a different color. An example of this is the LSMS survey of rural households in northeast China in 1995. Chinese characters do not have uppercase and lowercase, so two different fonts were used.

11. It is not necessary to convert quantities into standard units (for example, to convert bunches into kilos) to calculate farm

income, which was the purpose of the agriculture module in the Ghana LSMS. However, as is common with such rich data sets, analysts are using the data for other purposes as well, such as calculating the total quantities of various crops that were produced.

## References

- Ainsworth, Martha, and Jacques van der Gaag. 1988. *Guidelines for Adapting the LSMS Living Standards Questionnaires to Local Conditions*. Living Standards Measurement Study Working Paper 26. Washington, D.C.: World Bank.
- Babbie, Earl. 1990. *Survey Research Methods*. Belmont, Cal.: Wadsworth.
- Fink, Arlene. 1995. *The Survey Handbook*. Thousand Oaks, Cal.: Sage Publications.
- Fowler, Floyd. 1993. *Survey Research Methods*. Second ed. Newbury Park, Cal.: Sage Publications.
- Frey, James, and Sabine Mertens Oishi. 1995. *How to Conduct Interviews by Telephone and in Person*. Thousand Oaks, Cali: Sage Publications.
- Grosh, Margaret, and Juan Munoz. 1996. *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*. Living Standards Measurement Study Working Paper 126. Washington, D.C.: World Bank.
- Oliver, Raylynn. 1997. *Model Living Standards Measurement Study Survey Questionnaire for the Countries of the Former Soviet Union*. Living Standards Measurement Study Working Paper 130. Washington, D.C.: World Bank.
- Scott, Christopher, Martin Vaessen, Sidiki Coulibaly, and Jane Verrall. 1988. "Verbatim Questionnaires Versus Field Translation or Schedules: An Experimental Study." *International Statistical Review* 56 (3): 259–78.
- United Nations. 1985. "Development and Design of Survey Questionnaires." Department of Technical Cooperation for Development, National Household Survey Capability Programme, New York.

# **Part 2**

# **Core Modules**



# 4 Metadata—Information about Each Interview and Questionnaire

*Margaret Grosh and Juan Muñoz*

A survey data set should contain not only the answers given to the questions posed in the interviews but also some information about how the survey was conducted. This information is often called “metadata” or data about data.

Survey planners must make many different methodological decisions while they are designing their survey. Often, especially in the case of surveys carried out in developing countries, these decisions are made on the basis of anecdotal evidence, personal judgments, or “survey folklore” rather than on the basis of quantitative analysis of past experience. However, it is possible to gather evidence on various aspects of survey methodology to provide some guidance to the designers of future surveys.

Until now, there has been no systematic attempt in LSMS survey projects to collect and analyze metadata. It is ironic that LSMS surveys, which are designed to facilitate quantitative analysis of government policies and programs, have not been designed to apply the same types of quantitative analyses to themselves—that is, to facilitate analysis of survey methods. The aim of this chapter is to help to remedy this oversight by specifying what metadata should be collected in future LSMS surveys.

## **Issues**

Metadata can be useful in three different areas: for carrying out substantive analysis, for managing surveys, and for carrying out methodological research.

## ***Substantive Analysis***

Some metadata are required for analyzing survey data. Previous LSMS surveys have been reasonably effective collecting such data.

The most crucial metadata required for analytical purposes is information about the sample. Even simple analysis (such as means and cross-tabulations) requires correct sampling weights, which are calculated from information about the sample that must be collected during the planning and implementation of the survey.

Other metadata are useful for more specific analyses. Analysts often need to know the dates of the interviews in order to calculate important constructed variables. For example, the ages of respondents can be calculated by subtracting the respondents’ birthdates from the dates on which they were interviewed. It is also often important for analysts to have precise information on the ages of very young children so that they can use these data in conjunction with height or weight data to calculate the children’s nutritional status. The dates of the interviews are also important in adjusting any estimates for price differences that may have occurred due to inflation during the fieldwork period (especially when the fieldwork is spread over a year or, in high-inflation countries, even when the fieldwork is compressed into a few weeks) or between

two separate surveys when analysts are comparing data from the two surveys.

Occasionally analysts make more complex use of metadata. For example, they may conduct a sensitivity analysis to determine whether the patterns of answers given by replacement households or proxy respondents differ from the answers that would have been given by the intended respondents. If the answers are significantly different, the analysts may wish to omit the observations given by proxy respondents from their analyses or to adjust their calculations to correct for any biases introduced by the proxy respondents.

### *Managing the Survey*

It is extremely useful for survey managers to receive information on how the survey interviews are going while the survey is in progress. For example, survey managers may want to monitor how many interviews are being conducted in the various languages spoken in the country of the survey. If more interviews than expected are taking place in a particular language, the survey managers can arrange to have more questionnaire forms printed in that language. Similarly, if survey managers have ongoing information on how quickly the field teams are completing their work in a particular sampling unit, they can monitor whether the fieldwork is progressing either slower or faster than planned and can take steps to adjust the timetable or the budget accordingly. Also, if they know the rates of nonresponse or of proxy interviews for each interviewer during each month of fieldwork, they can arrange for the interviewers who are performing poorly to be supervised more closely, to receive remedial training, or, in exceptional cases, to be replaced.

### *Methodological Research*

Some insights gained during one survey may come too late to improve that particular survey but can be a valuable input into the design of subsequent surveys. Thus many of the methodological questions that metadata can help answer are the same as those involved in survey management but over a longer time frame. These questions include:

- How many more observations can be gained by having interviewers make a second, third, or fourth attempt at contacting a household where no one was home at the time of the interviewer's first visit?
- How long is the average interview?

- How much does the length of the interview affect the quality of the data?
- How do the characteristics of the household (such as size or economic activities) affect the number of times the interviewer must visit the household to complete the whole questionnaire?
- How do different characteristics of the interviewers (such as gender, age, or education) affect respondents' performance or responses?
- Into how many languages should the questionnaires be translated?
- How many languages should the interviewers be able to speak?

### **Data Needs**

To gather together a full set of metadata about a given survey, information must be gathered from several different sources. In this section, each of the main sources of metadata is discussed. In keeping with the theme of this book—the design of survey questionnaires—most emphasis is put on the data that are gathered in the questionnaire itself. This chapter concentrates on the issue of what data about the process of carrying out the survey should be collected and recorded rather than on explaining or giving recommendations about the survey process. The experience of and recommendations for LSMS surveys on sampling, fieldwork, data management, and project planning are the subject of a companion volume to this one (Grosh and Muñoz 1996). It should be noted that the recommendations made in this chapter about what types of metadata to gather will apply regardless of whether specific parts of the LSMS implementation logistics are adopted or not. The specific format may, of course, vary according to how the survey is implemented.

Metadata collected on the questionnaires should be typed into the computer and made available to researchers along with the survey data sets. In most previous surveys some metadata have been recorded on the paper questionnaires as an aid to field managers, but have not been entered into the computer. As a result, the metadata could not be used to help to analyze the substantive survey data or to rigorously evaluate the methods used in the survey.

### *The Household Questionnaire*

Some of the most important and useful metadata can be gathered by means of the survey's household questionnaire.

**IDENTIFYING INFORMATION FOR THE HOUSEHOLD.** The first and simplest piece of metadata on each questionnaire is its unique identification code.<sup>1</sup> No survey should omit this.

**SAMPLING INFORMATION FOR THE HOUSEHOLD.** For sampling weights to be calculated correctly, information from three separate sources must be brought together.

First, some information, such as the codes and names for each stage of the sampling process—the stratum, the number of the census block that serves as the primary sampling unit (PSU),<sup>2</sup> and the dwelling number—can usually be recorded on the questionnaire by the clerical staff who are preparing the survey before the interview is conducted.<sup>3</sup>

Second, as the interviews are conducted, further sampling information is collected for each household in the original sample. For example, it is occasionally impossible to interview a particular household in the sample. In these cases the interviewer should note the reason on the questionnaire (for example, the interviewer was unable to locate the building, the building was no longer being used as a dwelling, the residents were all absent for the duration of the fieldwork in the area, or the household members refused to participate in the survey). If other households are interviewed as “replacements” for these nonresponding households, the interviewer should note the fact that the interviewed household is a replacement.

Third, some information comes from the records of how the sample was drawn (see discussion below under “Sampling Records”).

**GEOGRAPHICAL INFORMATION FOR THE HOUSEHOLD.** The geographical information that should be collected for each household should specify the state (or department, province, or region) and the county (or municipality, district, or prefecture) within which the household is located. At least two, and in some countries three or four, levels of administrative unit may exist and, therefore, need to be recorded. This information can often be precoded into regions that are analytically meaningful, such as rural and urban areas or major agroclimatic zones such as mountains, plains, and coast. These codes can either be placed in the questionnaire or can come from the same administrative records from which the sampling weights are drawn.

It is also important to express geographical information not only in terms of administrative units but also in terms of geographical descriptors such as latitude and longitude (and even altitude). This information should be recorded for two units of observation—the primary sampling unit and the dwelling.

Gathering data on the location of the primary sampling unit, even within a few kilometers, will be helpful for analysts and is crucial both for managing the fieldwork and for merging the survey data, at least approximately, with other geographical information—for example, analyzing information on climate and soil quality in conjunction with data from the agriculture module. An increasing number of developing countries have data sets that include the location described in terms of latitude and longitude and the characteristics of government services such as clinics, schools, and post offices.

The geographical descriptors of the primary sampling unit in the LSMS survey can be drawn from the cartography done for the sampling. Alternatively, global positioning systems (also known as geopositioning systems) can be used during fieldwork to record the location of the community center and possibly of other important points of reference in the community (such as clinics, schools, the nearest hospital, the agricultural co-op, and the nearest paved road). The second option is likely to yield more precise measurements.

The locations of the dwellings of the sample households can also be recorded using global positioning systems, accurate to within a few meters. Knowing the exact location of households makes it possible for analysts to measure accurately the distance between each dwelling and the available services. It also allows them to merge the survey information more precisely with information containing geographical references from other sources. In addition, knowledge of a household’s exact location makes it easier for interviewers to find the dwelling again if the survey is repeated.<sup>4</sup>

Clearly, recording the location of the sampling unit is extremely useful and should be done without fail. However, recording the exact location of the dwelling raises a problem of confidentiality. To include the latitude and longitude of a household in the data files that are released to the public would mean that data users could, at least theoretically, locate specific households and connect their members with the information that they gave during the survey. This

would violate the confidentiality that interviewers promise respondents.

There are some potential solutions to this problem. One option would be to omit the latitude and longitude of the households from the files released to the public. This is parallel to the procedure often used for the names and addresses of the households, which are typically included on questionnaires but omitted from the publicly released data files. In this case the survey agency would have to calculate the distance variables before releasing the files to potential analysts. The implication of this would be that only the survey team would have the power to merge the household survey information with other geographical data. This might limit the number of such merges performed, since the survey agency would be less likely than a series of individual researchers working on many different issues to seek out complementary data sources on many different topics. An alternative option would be to give researchers access to the original geographical descriptor data files in order to create their own distance variables or to merge geographical data with other data sets, but to vet these individuals carefully to ensure that they did not make the information available to third parties or otherwise misuse it.

**RESPONDENTS FOR THE MODULES.** Interviewers should record which household member was the respondent for each module and each section. All of the draft questionnaire modules presented in this book specify a preferred respondent for each section. Broadly, they specify that every household member over a particular age (for example, ten years) should reply for himself or herself in modules that deal with individual attributes such as health, education, or labor activities. In the household enterprise and agriculture modules, the respondent for each enterprise or agricultural plot should be the household member best informed about that enterprise or plot. This system should yield the most accurate information. This system also has the advantage of spreading the burden of the interview among several respondents, thus reducing the chance that any one respondent will get tired.

Despite the survey teams' best efforts, there will inevitably be times when it will not be possible to interview every individual in person. In these cases interviewers will have to settle for proxy respondents. It is important that interviewers record when this happens, the reason why it happens, and the identity of

the proxy respondent, because analysts of the resulting data set may want to treat responses from proxies differently from responses given by preferred respondents. Also, survey managers may wish to monitor the rate of proxy respondents per interviewer as one way of judging the quality of each interviewer's work. Planners of new surveys may wish to know how much difference proxy responses make so they can know whether it is worth incurring the costs of making several visits to the household to reduce the number of proxy responses.

**INFORMATION ABOUT THE INTERVIEWS.** Interviewers should record the following pieces of information about the conduct of the individual interviews:

- *The length of time taken to interview each person.* Recording this information is essential so survey managers can know how onerous a burden is put on respondents during the survey. Interviews can be split into more than one session. Thus the basic unit of observation should be the "sitting" or the length of time a given respondent is interviewed on one occasion. The total interview time for that person is the sum of the time spent in all of his or her sittings. The total interview time for the household is the sum of all sittings by all individuals in the household. In some surveys it may also be useful to know how long it took in each household to complete a specific module, especially for modules that are new, experimental, sensitive, or very lengthy.
- *How many times an interviewer must visit a household to obtain all of the required information.* Survey managers and planners want to know whether it is worthwhile in terms of time and costs to make third or fourth visits to households in order to get more modules filled out or to minimize proxy responses. Therefore, interviewers should record the number of visits they make to each household and the number of person-modules<sup>5</sup> completed on each visit.
- *The date each module was administered to each household.* This information must be recorded: to help analysts calculate precisely the ages of the respondents (especially children) based on their birth dates and interview dates; to help analysts adjust any monetary values expressed in the responses for inflation; to analyze any seasonality issues; and to find out how long a period elapsed during the reference period for questions that include a time-frame (for example, the answer to "How much have

you spent on [..item x..] since my last visit?” can only be calculated on a 12-month basis if the number of days between the two interviews is known).

- *The language in which the interview is conducted.* It is sometimes prohibitively expensive or complicated to translate the survey questionnaire into all of the languages spoken in a country. In these situations, when the respondent does not know any of the languages in which the questionnaire is written, the interviewer has two options. First, he or she can conduct the interview in a language other than that in which the questionnaire is written (for example, the LSMS questionnaire for Ghana was written only in English, but some interviews took place in Akan). Second, if the interviewer and the respondent speak no language in common, the interview may be conducted through an interpreter. This is often a person who resides in the household or primary sampling unit where the interview is conducted. If an interpreter is used, the interviewer should record this. Since errors are inevitably introduced in the process of translation and interpretation (see Scott and others 1988), it is important to monitor the overall rates of these occurrences and to let analysts know for which interviews these conditions held. It is usually sufficient to record this information for each individual rather than for each person-module, since individual respondents tend to use the same language to respond to all their modules. For example, if an elderly Bolivian woman responds to the questions in the health module in Quechua rather than in Spanish, she will probably also respond to the employment module in Quechua rather than Spanish. Her son, however, might be interviewed in Spanish.
- *Whether other people are present during an interview and what the relationship is between the respondent and the observer.* Respondents are more likely to reveal full and accurate details of sensitive information on such topics as income, savings, and sexual behavior if the interview is conducted in private rather than in the presence of others. For example, a wife may not want her husband to know how much money she has earned or saved from her farm plot. A youth may not want his or her parents to know about his or her sexual encounters or contraceptive practices (see, for example, Hoyt and Chaloupka 1994 on the willingness of young people to report their use of drugs in the presence of others). This general prin-

ciple applies to all modules but is presumably more relevant in some than in others. Survey managers may wish to monitor the extent to which other people are present during interviews because if this occurs frequently in the case of some interviewers, it may be a sign that they lack certain skills or are not being sufficiently diligent in following the fieldwork guidelines.

- *Notes in the margins of the questionnaire or in special sections provided for notes.* These can be useful during the process of cleaning the data files. For example, on one occasion, an interviewer noted that the lady being interviewed attributed her substantial income not to the employment or assets she was asked about in the questionnaire, but to her two very generous lovers. On another occasion, an interviewer wrote “respondent dead drunk, answers probably not true.” In both these cases the comment provided was very useful. In the first case the responses to the questionnaire were accurate but would probably have been discarded during the data cleaning process had it not been for the interviewer’s comment. In the second case the information on the questionnaire was probably inaccurate, but it might have been used had the interviewer not made that note.

**INFORMATION ABOUT THE SURVEY TEAM.** The code number of the field staff (the interviewer, supervisor, anthropometrist, and data entry operator) associated with each household should be noted on the questionnaire, as each of these individuals deal with that household. This makes it possible for survey managers to monitor the quality of their survey teams. It also allows analysts to look out for “interviewer effects” that might complicate or distort analysis. Planners of subsequent surveys and methodologists often wish to analyze how various features of the interview (such as completion rates, replacement rates, nonresponses by item, and proxy respondents) vary by the characteristics of the field staff.<sup>6</sup> In order for them to be able to do the fullest possible analysis, it is necessary to prepare a separate file (taking the information from the administrative records of the agency conducting the survey) that lists all of the characteristics of the field staff, including gender, education, years of experience in the job, ethnicity, age, and languages spoken. A format for recording information from the administrative records is introduced in the third section of this chapter (and provided in Volume 3).

*The Community (and Facility) Questionnaire(s)*

The same list of metadata that needs to be collected on the household questionnaire also needs to be collected on the community and facility questionnaire(s), with only two variations.

The first exception is the geographical location information. While it may be acceptable not to record the latitude and longitude of households, this should be done as a matter of course for communities (and facilities), as discussed above.

The second variation involves the privacy of an interview. In household interviews the interviewer should try to ensure that each respondent can give his or her responses in private. In contrast, when community questionnaires are fielded, interviewers explicitly seek out groups of people to be interviewed together. In many cases this means that answers given will convey some kind of consensus or common view in the community. In other cases group interviews yield a full range of different answers. For example, in answer to a question about what sources of health care are available to the community, a group interview may elicit a longer list of possible places to go for health care than would an interview with a single respondent. The interviewer should record information about the people who contribute to the group interviews, including their names, genders, positions of leadership in the community, approximate ages, and, possibly, ethnic groups.

However, there is a category of metadata that can only be collected within the community questionnaire: the costs and logistics of the survey by primary sampling unit. Many of the items that influence total survey costs are specific to the primary sampling unit and not to the household, so the primary sampling unit/community is the appropriate unit of observation. Having information on such questions as how many nights the interviewers had to stay in the primary sampling unit overnight, whether they stayed in hotels or in households, the number of miles the vehicles traveled between and within primary sampling units, and whether electricity was available for entering the data into computers helps planners of subsequent surveys draw up accurate budgets and work programs.

*Sampling Records*

Some important sampling information needed to calculate the final sampling weights has to come from the

sample frame and from the records that document how the first stage of the sample was drawn. Thus it is important to ensure that the administrative, geographical, and sampling terminology that applies to each primary sampling unit, the probability of that particular unit being selected into the sample, and the measure of the size used in drawing the first-stage sample are recorded. How these should be recorded is discussed in the third and fourth sections of this chapter.

*Administrative Records*

Many important metadata are contained in administrative records rather than in the survey questionnaire, including data relating to the survey's field staff. Information on the interviewers is especially important since these are the staff who have the most contact with the sample households. However, the same information should be recorded for all team members, including supervisors, data entry operators, and anthropometrists as well as interviewers. An important question in survey work is whether the characteristics of the interviewer affect how the respondents answer the survey questions. Therefore, it is important to collect information on the interviewers' characteristics—at minimum their age, gender, and education. In places where many of the interviewers have more than a secondary education, it is useful to note in which discipline they are trained. The amount of survey experience each interviewer has had should also be noted, as well as whether they are permanent staff of the statistical agency or are short-term contractors. Often, the race or ethnic group to which interviewers belong will also be pertinent. A sample form for collecting this information is introduced in the third section of this chapter (and provided in Volume 3).

Cost information is another important aspect of metadata that can be gleaned from administrative records. Planners of subsequent surveys are likely to want to see basic budget information about the survey, such as how many staff worked at what salary for what period of time,<sup>7</sup> how many and what kinds of durable equipment were required (for example, vehicles and computers), and how many and what kinds of consumables were used (for example, paper, pencils, diskettes, printer cartridges, and field kits for interviewers). No form is given for this in Volume 3 because the style of every survey's budget is usually dictated by the requirements of the various agencies that finance the survey. Those who analyze metadata

from a survey after it has been completed should use the final budget figures rather than the budget devised when the survey was originally being planned.

#### *Data Entry Program*

The data entry program can record various statistics about each data entry session for each household (for example, the date on which the entries were made, the length of time it took to enter the data, the operator's identification code, and the number of entries that showed errors when range or consistency checks were performed). This information has hardly ever been analyzed in the case of previous LSMS surveys, but it could reveal a lot about how the fieldwork was actually carried out. For example, it could help answer questions such as: In how many cases were the data from the first part of the questionnaire really entered before the second set of interviews took place? How long did it take to enter the data from a questionnaire? How often did errors in skip patterns occur?

#### *Special Experiments*

To analyze some issues a special experiment may be required. Consider, for example, the issue of whether the codes for "yes" and "no" should be 1 and 2 (as in most previous LSMS surveys), 1 and 3 (as in RAND's Family Life Surveys), or 1 and 6 (as in some other surveys). The idea behind using 1 and 3 or 1 and 6 is that the data entry operator is likely to make fewer mistakes because the keys for the two different numbers are further apart on the keyboard than 1 and 2. However, since the normal administration of surveys includes checks for data consistency and/or for double-blind entry of data and since the entered data are proofread and checked against the original questionnaire pages, most (if not all) of these typographical errors should be corrected before the data set is produced. Thus a special experiment that included data entry without all of the normal quality controls but with a special assessment of accuracy would be required in order to definitively answer the question of what are the best codes to use for "yes" and "no." Since these experiments are not commonly done as part of general survey implementation, no more will be said about them here. They have been mentioned only to remind the reader that the metadata that can reasonably be gathered as part of implementing a survey, while useful, will not answer all questions of survey methodology.

#### **Draft Module**

Most of the metadata from the household questionnaire can be gathered in a special "metadata module" of the questionnaire. A draft metadata module is presented in Volume 3. The metadata module consists of three parts: the household identification page, the record of visits and interviews, and the comments page. Alternatively, some of the information can be collected on the pages of the topic-specific modules using an alternative design of the questionnaire page, examples of which are presented in Volume 3. Forms are also included for collecting important metadata from the community questionnaire and from sampling and personnel records. Detailed notes on how to use

#### **Box 4.1 Cautionary Advice**

- *How much of the draft module is new and unproven?* The household identification page is similar to those used in many previous LSMS surveys. The record of visits and interviews changes the unit of observation from the whole questionnaire to the "sitting." This change is untested, but most of the information collected in the grid was previously collected for the household as a whole. Neither the field staff's comments nor the characteristics of the survey staff have been incorporated into the electronic files in previous LSMS surveys, but these have both been done in other surveys with stronger traditions of collecting metadata than the LSMS. The sheet for survey costs and conditions specific to the primary sampling unit is new. The information contained in the sampling spreadsheet has always been collected in one way or another but usually not so tidily packaged until late in the data-cleaning phase.
- *How well has the module worked in the past?* The household identification page has worked quite well in previous LSMS surveys. The household-level record of visits and interviews may not have been filled out as rigorously as it could be and has rarely been used in analysis, probably because it was evident to both interviewers and analysts that the unit of observation used in the past did not make sense.
- *What parts of the module most need to be customized?* Survey planners will have to choose what amount of metadata to collect. Thereafter, the main customization required will be the nomenclature for the various sampling terms, administrative terms, ethnicities, and languages. Also, if cost data are collected in the community questionnaire, this will need to be customized to fit the fieldwork plan.

these forms are provided given in the fourth section of this chapter.

### *The Household Metadata Module*

The first part of the module consists of the household identification. This should be filled out for every household in the sample as originally drawn, regardless of whether the household is interviewed. This part contains the sampling information (including whether the household is a replacement and, if the interview did not take place, why not), the geographical information (including the address of the dwelling), minimal information about the head of the household, and the codes for the field staff.

The second part of the module is a grid for recording a summary of visits and interviews (otherwise known as the sittings grid). In this grid, the unit of observation is the “sitting,” which is the length of time any one respondent is interviewed at one time. On a single visit to a household, the interviewer may conduct more than one sitting if she or he interviews more than one person. The grid records important information about the sitting—how long it lasted, whether the person interviewed was responding for himself or herself or as a proxy, the language used, whether or not others were present, and the interviewer’s impression of the quality of the data. A space is provided on this form for matching specific comments that appear on the next page (the field staff comments page) to the sitting to which they pertain.

An alternative to the summary of visits and interviews page would be for the interviewer to record the respondent’s name and ID code number, the language used, and whether others were present during the interview somewhere within the topic modules. This would be straightforward as the interviewer can observe all of these details as a matter of course rather than needing to ask questions about them. In some modules it may even be feasible to add some items that the interviewer would have to take action to discover, such as the start and end time of each module (by looking at his or her watch) or the reason why a proxy respondent had to be interviewed. However, doing this for each and every module would probably disrupt the flow of the interview. If interviewer checked his or her watch at the beginning and end of each module, the respondent would probably either feel rushed or become impatient to finish the interview.

The third part of the household metadata module is the comments page. Interviewers should be asked either to write all of their comments here during the interview or to transfer the comments that they wrote down elsewhere to this space after the interview. Putting all of the comments together in one place makes it much easier for the data entry operator to enter them in a text file. The system for numbering the comments and for cross-referencing them to the sittings grid, section, and question to which they pertain is designed to make it easy for the analyst to locate appropriate comments.

There are two circumstances in which survey designers may wish to arrange their survey so that the metadata are collected in the topic modules. The first circumstance is when designers do not need to collect a full set of metadata but simply want to collect metadata on modules that they think contain the most sensitive topics or on modules that are experimental or may not be as well designed as the others. The second circumstance is when designers want to collect not only the metadata in the sittings grid but also some extra metadata on a few modules. In this case they should include the sittings grid in the questionnaire and leave space for collecting metadata in the topic modules as well. For example, survey planners may want to collect metadata on the average time it takes to complete a specific module. In this case they should leave spaces for the interviewer to record the time at the beginning and end of the module in the module itself. The information on the sittings grid will not be sufficient because several modules are usually filled out during a single sitting.

The form following the comments page shows how information on the respondent and on the length of the interview can be added to the food page of the consumption module. In reality it is probably of more interest to analysts to know how long it took to complete the whole consumption module than to complete each part of it. Therefore, the blanks for recording the start and stop times might begin on page 1 of the consumption module and end several pages later. The dotted line indicates that questions have been omitted to show the beginning and end of a multipage module on a single sheet. Another form (in this case, a section of the employment module) shows how information on the interview can be added to an individual-level module. Again, the time recorded for the “length of interview” should probably cover the

whole employment module rather than just one part of it.

The reader should note that the forms presented here differ from past LSMS practice in two ways. First, the sittings grid is more detailed. Previous LSMS surveys were organized as follows. Interviewers made a first visit to all sample households in a given primary sampling unit, at which time they administered the first half of the questionnaire (roughly speaking, the individual-specific modules). This was followed by a period when the data from these first interviews were entered in the computer, while the interviewers visited households in a different primary sampling unit. Then, the interviewers returned to the sample households in the first primary sampling unit to conduct a second “interview” to administer the remainder of the questionnaire. Metadata were, therefore, collected for each “interview.” However, the drawback to using the “interview” as the defined unit of measurement was that, within each “interview,” interviewers were asked to conduct “mini-interviews” with each household member over the age of ten. Accomplishing this often required more than one visit to a household. Thus the format given for collecting information by “interview” did not correspond with actual field practice. This may be one of the reasons why little use has been made of the metadata information collected.

The second way in which the draft module differs from those included in previous LSMS surveys is that it makes it possible to enter interviewers’ comments in the electronic data files. In previous surveys interviewers have always written marginal comments on the questionnaires, but these comments have not been included in the electronic data sets. Thus analysts have been unable to consult these comments, even when they may have been necessary to elucidate the data.

Thus, in the forms proposed here, the unit of observation has been changed to reflect the actual practice in the field. However, it is important to bear in mind the fact that the sittings grid and comment system presented here have not been field tested. It will be important to conduct careful field tests of both instruments before using them in a specific country setting, to ensure that both interviewers and their supervisors are properly trained in how to use them.

As with all the other modules, the metadata module will need to be customized to each country context. For example, on the household identification

page, the nomenclature and number of codes will have to be changed to fit the circumstances in the country where the survey is to be conducted. Also, in the sittings grid, the list of languages in Question 7 and the list of modules in Question 10 should be changed to match the survey being conducted.

#### *Form for Collecting Information on Survey Costs*

The purpose of this form is to collect information on the cost of the survey’s fieldwork, thereby helping the planners of similar surveys in the future. Including this form as a page in the community questionnaire would simplify the administration of the survey and make it less likely that the form would be lost or not filled out. However, since the information collected by this form has a purpose so different from the information collected in the rest of the community questionnaire, the form has been discussed and included here instead of in Chapter 13.

#### *Form for Recording Information from Sampling Records*

This form is for recording all pertinent sampling information. Some of this information will be generated in the course of drawing up the first stage of the sample, some will come from the results of the listing operation, and some will come from the results of the survey itself. Also, some will come from calculations based on the raw information from the above three sources. Details of this process are given in the fourth section of the chapter.

#### *Form for Recording Information on Personnel*

##### *Characteristics*

This form is for recording all pertinent information on the field staff. Some of this information may be available from the personnel records of the statistical agency, although these records often omit details—such as the languages spoken by the staff member or the staff member’s ethnicity—that can be important to analysts studying interviewer effects. One way to gather all the necessary information might be to do so at a training session that brings all staff together in the same place at the same time.

#### **Explanatory Notes**

This section of the chapter provides some explanatory notes to the questionnaire forms given in the previous section.

*Notes on the Household Identification Page*

The household identification page is frequently referred to as the “cover page” and often is, in fact, the cover page of the household questionnaire. This is not desirable, since the covers of questionnaires are subject to a lot of physical strain as they are handled during the course of fieldwork and transportation. They tend to come loose, and if the information contained on this page is separated from the rest of the questionnaire, the rest of the questionnaire will be useless as it will not be possible to assign the proper sampling probabilities to the information collected. Moreover, if household-specific information, no matter how innocuous, can be read on the front of the questionnaire, this may make respondents skeptical of the interviewers’ assurance that the information collected in the survey interviews will be confidential.

Thus the cover page should consist of a page of stiffer material emblazoned with the logo of the statistical agency or survey. The only piece of household-specific information that should appear on the cover page should be the household’s unique identification code. This code should also be recorded on this form. However, to ensure that there are no double-recording errors, it may be better to record the code on the household identification page but to arrange for it to appear through a cut-out window on the cover page.

The household identification numbers, the name of the head of household, the household’s address, and, sometimes, the household’s geographical codes can be filled out on the household identification and control information page either in the survey office or by the supervisor before the interviewer takes the questionnaire to the household for the first time.

There is room on this form to note the number of attempts that the interviewer makes to contact the household. These include the interviewer’s attempts to meet household members and persuade them to be interviewed, as it may take several visits to the dwelling before the interviewer finds someone home. If a contact is made, either a date for an interview may be made or a sitting may begin immediately. Only once the sittings begin does the interviewer need to begin to fill in the sittings grid. (The total number of trips that an interviewer makes to a household is the sum of the number of contacts from the contact list and the number of different visits to the household derived from the sittings grid.)

The labels for the administrative and sampling codes should be changed in accordance with the local nomenclature.

If the household has a telephone, the interviewer can enter the telephone number of the dwelling beside the address. It may later be possible for the survey supervisor to make phone calls to such households to double-check that the interview really took place and that key pieces of information were collected correctly or to clear up any confusion that may have been detected by the data entry software.

The household’s ethnic group and religion are sometimes noted on the identification page. These are, of course, individual-specific traits; different household members may have different ethnic origins or religious affiliations. Thus, conceptually, this information should be on the roster page of the household questionnaire. However, in many countries, it is deemed impolite to ask direct questions about such matters, and often the interviewer can easily observe these characteristics without asking any questions. In such cases the information may be placed on the household identification page. The example page in Volume 3 is formatted this way to remind survey designers of this option. There should always be an “other” category in the list of ethnicities.

*Notes on the Summary of Visits and Interviews*

Most of the information on the grid is straightforward. Notes are provided here only as necessary.

**Q4.** It is important to equip interviewers with watches if it is likely that some interviewers will not have them.

**Q7.** The list of languages should be customized to fit the languages spoken in the country. It should include the languages into which the questionnaire was translated, other languages spoken by significant subgroups of the population, and an “other” category.

**Q9.** It is possible to include more details here. The ID code of the household members who are present can be recorded. Also, the list of nonmembers can be coded either by demographic characteristics (such as age or gender) or by the type of social relationship that the respondent has with the nonmember (such as kin, neighbor, or village elder).

**Q10.** Below the main grid is a space for the interviewer to mark whether a module is required for any

particular household. If a household has no household enterprise, that box can be marked off on the row below the grid. This will make it easier for the interviewer to tell whether or not they are finished with a household by looking at the sittings grid rather than having to look through the whole questionnaire.

**Q11.** The interviewer should provide his or her assessment of the reliability of the information given in the interview in column 11. The assessment is qualitative but is based on detailed information that only the interviewer can have. The issue of what constitutes reliable and unreliable information should be discussed when interviewers are trained.

**Q12.** In column 12 the interviewer should note the number of textual comments he or she may have written on the comments page that pertain to a particular sitting.

#### *Notes on the Characteristics of the Survey in the Primary Sampling Units Form*

In cases in which one community questionnaire is filled out for each primary sampling unit, it is best to integrate this form (which gathers information on costs specific to the primary sampling unit) into the community questionnaire. If several community questionnaires are filled out for each primary sampling unit, this sheet can be treated as a separate form. However, such separate forms are less likely to be filled out thoroughly or to be key-entered into computer files, since they are often not viewed by the survey team as one of the major tools of the survey. Thus, wherever possible, questions on costs specific to the primary sampling unit should be included in the community questionnaire rather than collected on a separate form.

Questions 1–4 may have to be modified to fit the fieldwork plan. According to the wording used here, for example, the teams are assumed to travel from one primary sampling unit directly to another without visiting the headquarters of the survey team in between. If this is not how the fieldwork is organized, these questions should be changed accordingly.

Similarly, the version of this form given in the third section of this chapter assumes that each team has the use of its own vehicle to ferry the interviewers to the various places within the primary sampling unit that they need to visit. If the interviewers have to travel by other means (such as by public transport, boat,

motorcycle, bicycle, horseback, or on foot), the questions should be modified to fit these circumstances. This may need to be done carefully as different means may be used in different parts of the country, but the form must be relevant to all of these circumstances.

#### *Notes on the Sampling Spreadsheet*

The sampling spreadsheet is a byproduct of the first stage of sampling. (For details on how this works for LSMS surveys see Grosh and Muñoz 1996.) Each line in the spreadsheet corresponds to one primary sampling unit selected from the whole sample frame to serve as part of the sample. Some of the columns in the spreadsheet contain data straight from the sample frame whereas others are computed within the spreadsheet or come from other sources, as explained below.

Column A (PSU No.) is a short identifier for the primary sampling unit for the purposes of a specific household survey. It is convenient to use mnemonic, three-digit identifiers such as these rather than the long chain of “geocodes” that statistical agencies sometimes use to identify the units in the sample frame. In this case the identifier is a simple serial number within each stratum (for example, 101 to 125 for the 25 sampling units selected in stratum 1 and so forth), but it would be equally appropriate to use a straight serial number from 1 onwards to identify the primary sampling units across all strata.

Columns B and C are for the stratum codes and names.

Columns D to L are for the geocodes used by the statistical agency of the country surveyed. The situation depicted in the spreadsheet may be oversimplified because: sometimes the names of the hierarchical subdivisions differ between urban and rural areas; in some countries intermediate categories between “urban” and “rural” are used; and although larger geographic subdivisions usually have names and codes whereas smaller ones have only codes, the specific definitions of “larger” and “smaller” are country-specific. In the spreadsheet presented in Volume 3, space is given for the codes and names of the region and county; however, if smaller geographical subdivisions have names, these should also appear on the spreadsheet.

Columns M, N, and O are for recording the longitude, latitude, and altitude from the approximate center of the primary sampling unit as taken from the sampling cartography. This method is less precise than using global positioning systems in the field, but it still enables

some approximate merging of these data with other sources of geographical data. These columns can be deleted if global positioning systems are used in the field.

Columns P to T are for recording the details of the first sampling stage, assuming that it is done with probability proportionate to size—which is almost always the case. In practice, columns P and Q would be headed “Number of households in the stratum” and “Number of households in the PSU” (or “Number of dwellings” or “Population” or whatever actual measure of size was used in the selection with probability proportionate to size). Note that column P (the size of the stratum) is a constant for all primary sampling units within a stratum. Column R (the number of primary sampling units selected in the stratum) is also a constant within the stratum.

Columns S and T are formulae.  $S=Q(R/P)$  is the probability of selecting the PSU in the first sampling stage, while  $T=1/S$  is the contribution of the first sampling stage to the raising factor of households in the PSU (in other words, how many households in the total population the sampled households represent).

Columns U to Y record the details of the second sampling stage. U is the number of households actually recorded by the listing operation in the PSU.<sup>8</sup> It should be completed by survey managers as the results of the listing operation come in to the survey headquarters.

If the measure of size used in the first stage is the number of households, it may be useful to compute an auxiliary column with the ratio  $U/Q$  and use it to monitor the quality of the listing operation. If both the listing operation and the census that originated the sample frame are carefully carried out,  $U/Q$  should be a little larger than 1.0 almost everywhere (except in areas depopulated by natural disaster, war, or other unusual events). On average,  $U/Q$  should approximate the expected population growth rate for the country. If it measures less than 1.0 in a lot of primary sampling units, this may be the result of sloppy work during the listing operation.

The entries in columns V and W should come from the survey’s data set. This underscores the importance of leaving room to record refusals and other kinds of nonresponse in the household identification page of the household questionnaire and of actually completing and key-entering the data from the household identification pages of all of the questionnaires, even those from nonresponding households. It may be useful to split column W into the various kinds of

nonresponse recorded on the questionnaire. In most previous LSMS surveys the sample was designed so that the same number of households in each primary sampling unit was interviewed, which would thus make column V almost constant. In the example given in the spreadsheet (in Volume 3) the sample was designed so that 12 households per primary sampling unit were selected in the second stage, although this number could be different (for example, 16).

Columns X and Y are formulae.  $X=V/U$  is the probability of selecting each particular household in that primary sampling unit in the second sampling stage, while  $Y=1/X$  is the contribution of the second sampling stage to the raising factor of households in the primary sampling unit.

Columns Z and AA are also formulae.  $Z=SX$  gives the final selection probabilities, while  $AA=TY$  gives the raising factors for households in the primary sampling unit.

Columns AC, AD, and AE are the distribution of the sample across teams and throughout the year. AC records the time period during which the primary sampling unit should have been visited according to the original schedule. AD is the actual date of the first interview in the primary sampling unit, and AE is the date of the last interview in the area. Analysts will want to verify that the pattern of actual interviews does not depart too much from the prescribed pattern in cases in which the fieldwork is meant to be spread evenly over the year and across the country to eliminate seasonality effects. Survey administrators will want to note whether the expected pace of fieldwork was maintained in the field.

#### *Notes on the Staff Information Form*

Information about field staff is essential for analyzing how interviewers affect respondents’ answers. Therefore, all field staff positions should be included on the staff information sheet. It may also be worth including nonfield staff here if the information is likely to have other administrative uses.

**COLUMN A.** Each staff member should have a unique staff number, which might be a number used in personnel records. Alternatively, a number may be assigned for the purpose of this survey, starting with 001 for the first staff member and continuing upward as more staff are hired. If a staff member leaves, his or her replacement should be assigned a new staff num-

ber rather than being assigned the number of the staff member who has departed; reusing numbers might be confusing for analysts.

**COLUMN F.** During fieldwork, staff may do jobs that are not their primary function. For example, anthropometrists may pitch in to help fill out community questionnaires or supervisors may conduct an interview or two to keep the team on schedule. By comparing the roles listed here with who actually fills out the various parts of the sittings grid, it is possible to determine how much of this substitution occurs. Evaluating this after the fact may yield conclusions that would be useful inputs for designers of future surveys regarding workloads and training needs.

**COLUMN G.** The levels of education and the terminology used will need to be tailored to the country of the survey. If interviewers in a given country have commonly had post-secondary training, for example, it may be worth noting the main disciplines from which interviewers are drawn.

**COLUMN I.** The terminology for contractual status should also be tailored to the country studied. The main distinction to be made is between the permanent staff of the survey agency and workers who are contracted only for the duration of one survey.

**COLUMN J.** The main ethnic groups prevalent in the country should be used.

## Notes

The authors would like to express their gratitude to Martha Ainsworth, Didier Blazeau, and Teresa Parsley Edwards for the assistance provided.

1. Occasionally a household has so many members that two questionnaire forms are required to record the necessary information about all of the household members. In this case the two questionnaire forms should have the same household ID code, but the identifier codes for the people within them should be different; household members 1–12 would be in the first booklet, while members 13–24 would be in the second booklet.

2. If the sample is drawn in three stages rather than two, the information is required on both primary and secondary sampling units. This chapter assumes that the sample is drawn in two stages, but the examples can easily be extended to cover an additional stage of sampling.

3. This information can be recorded on the questionnaire in several ways. Ideally it would be done by an automated procedure in the central office to reduce transcription errors. For example, labels might be printed directly from the sampling files and then peeled and placed on the questionnaire. Of course, this assumes that both electronic files and facilities for printing labels are available, which is not always the case. Thus either the information from the sampling files can be transferred to the questionnaires in the head office before the questionnaires are given to the interview teams or the supervisors can transfer the information from the “assignment roster” onto the questionnaires. The codes for the region (rural/urban or agroclimatic zones) can also be placed on the questionnaire or drawn from administrative files.

4. Moreover, if geopositioning systems are used during the listing operation prior to the interviewing phase, it may become easier to locate the sampled households.

5. A person-module would be, for example, the administration of the employment module to one person. A single sitting is likely to cover the administration of several modules to one person, and a single visit by the interviewer to the household is likely to involve sittings with different individuals.

6. It is also interesting to study how these vary by the characteristics of the population. Moreover, since the assignment of interviewers is rarely totally random (for example, interviewers posted in rural areas will interview more farmers, while interviewers that speak an ethnic language will interview more respondents than average who speak only that language), it is important to control for the characteristics of the respondents when studying interviewer effects. However, no extra data about respondents are likely to be required since the typical control variables (such as region, age, sex, ethnicity, and occupation) are gathered in the course of administering the other modules.

7. It is assumed here that interviewers are paid fixed monthly salaries (which has usually been the case in previous LSMS surveys). If interviewers are paid on an interview-by-interview basis or if they receive bonuses for completing interviews fully or before some target date, that information should also be recorded.

8. In the listing operation, field staff should conduct an inventory of all of the dwellings within the boundary of each sampling unit. The resulting list of dwellings is the basis for selecting dwellings within the sampling unit.

## References

- Grosh, Margaret, and Juan Muñoz. 1996. *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*. Living Standards Measurement Study Working Paper 126. Washington, D.C.: World Bank.
- Hoyt, Gail Mitchell, and Frank J. Chaloupka. 1994. “Effect of Survey Conditions on Self-reported Substance Use.” *Contemporary Economic Policy* 12: 109–21.

Scott, Christopher, Martin Vaessen, Sidiki Coulibaly, and Jane Verrall. 1988. "Verbatim Questionnaires Versus Field

Translation or Schedules: An Experimental Study." *International Statistical Review* 46 (3): 259–78.

# 5

## Consumption

*Angus Deaton and Margaret Grosh*

The living standards surveys aim both to measure and to understand living standards. Much of the focus is on poverty or deprivation—the lack of adequate living standards. Standard economic measures of deprivation are concerned with the lack of goods or the lack of resources—income or assets—with which to obtain goods. But it is always important to keep in mind that many of the most important aspects of deprivation go beyond purely material deprivation. Deprivation of health, deprivation of education, deprivation of freedom from crime, and deprivation of political liberty are all important—often more important than deprivation of material living standards. The role of development in freeing people from deprivation in a wide sense has been forcefully argued by Amartya Sen. (See Sen 1999 for a recent and comprehensive account.) Data from the living standards surveys, particularly from the modules on health and education, often help us take a broad view of poverty. Other important aspects of living standards, such as life expectancy, infant mortality, and the threat of crime, must be examined using other types of data. Nevertheless, measuring the material basis of living standards will always play an important role in the assessment of levels of living. This chapter explores ways to collect data for a consumption-based measure.

The measurement of consumption has been a central objective of the LSMS program since the program's inception in 1980 and has remained so throughout some 50 surveys, all of which have been used to document living standards and poverty. Although the program has always acknowledged that living standards have many dimensions and has taken care to measure them in its surveys, the narrowly economic aspect of living standards in the program title was taken to mean not income, as had been the case in many previous surveys, but consumption. This consumption focus differentiated the LSMS surveys from some surveys in developing countries that immediately preceded the

establishment of the LSMS, such as RAND's Malaysian Family Life Survey. However, many earlier surveys, including the Indian National Sample Survey, had long used per capita household expenditure as the measure of living standards and the basis for measuring poverty.

The basic ideas underlying the measurement of consumption are straightforward. Nevertheless, there are various practical complexities, many of which are discussed in this chapter. Although income and wealth are what enable people to obtain goods and services, it is those goods and services themselves that directly generate economic well-being. The consumption

module of the LSMS survey is designed to measure the consumption of these items in some detail and in the aggregate (with the aggregate being the total value of consumption at suitable prices). At its simplest, the module collects data on how much people spend on various goods and services. How best to gather such information and in how much detail, how to deal with goods that are not obtained through the market, and how to obtain accurate data on prices are among the topics discussed in this chapter.

The LSMS surveys differ from many other household surveys in that their primary concern is not the estimation of means or totals. The most important concern of the LSMS surveys is documenting the distribution of living standards—measuring poverty (often, but not exclusively, the fraction of the population in the left tail of the distribution) and, to a lesser extent, inequality. The LSMS data are also used to illuminate a wide range of policy issues from descriptive tabulations to econometric modeling.

This emphasis on poverty and distribution must constantly be kept in mind because it has implications for the design of any LSMS survey. A survey that yields accurate estimates of average levels of income or consumption may nevertheless do a poor job of documenting income and consumption among the poor or of estimating the inequality of incomes. For example, if people have difficulty remembering high-frequency purchases (for example, food) after a day or two, asking respondents about the purchases they made on the previous day will yield more accurate data than asking them about the purchases they have made over the previous week or month. If the main concern is to estimate mean expenditure for the population, it may be sufficient to collect data on the average of the previous day's consumption since this figure would include all purchases both for those who purchased nothing the previous day and for those who purchased several days' supply. In contrast, the average of the previous week or month's purchases will be biased downward if the longer recall period implies that some purchases will be forgotten and thus not reported. However, for measuring poverty, the previous day's measure will not be sufficient because all those who did not purchase anything would be counted as poor. This means that to measure poverty it might be better to ask people about purchases they made over the previous week or month, despite the resulting downward bias.

Much of the literature on the design of surveys in general is concerned with how best to estimate means and totals and can be seriously misleading when applied to LSMS-type surveys, which have different concerns. This is particularly true of consumer expenditure surveys, which are most often designed to collect weights for consumer price indexes.

The first section of this chapter briefly reviews the arguments for using consumption rather than income both to measure living standards and to measure poverty and inequality. The first section goes on to discuss the principal uses to which consumption data have been put, including the documentation of living standards (still the central aim of LSMS surveys) and the illumination of a number of other important policy issues. Finally, the first section reviews some of the experience of more than 10 years of LSMS surveys in collecting consumption data.

The second section discusses the data needed to construct a consumption-based measure of living standards and reviews the design issues that affect both the cost of collecting data and the data's eventual accuracy. The third section presents a draft consumption module. The fourth section provides explanatory notes regarding the draft module.

## Policy Issues

In many cases consumption data are better than income data for measuring living standards. In addition, consumption data have a number of important analytical uses in their own right.

### *Why Use Consumption to Measure Living Standards?*

Although the LSMS surveys, like many surveys in developing countries, give primary emphasis to consumption rather than income, a considerable number of other surveys concerned with well-being do not attempt to collect consumption data. Many of these surveys are in industrialized countries, but the income focus is also standard in most surveys in Latin America. There are both theoretical and practical considerations that affect the choice of income or consumption, and the balance in favor of one or the other may be different in different circumstances. Thus it is useful to start by rehearsing the main arguments for and against each measure.

**THEORETICAL ISSUES AND IMPLICATIONS FOR MEASUREMENT.** Income and consumption are different

concepts, not just two different ways of measuring the same concept. Some economists prefer income as a measure of living standards because they follow a “rights” approach. According to this approach, income, together with assets, measures the potential claims on the economy of a person or family. Other economists prefer to use consumption because they consider the level of living a measure of economic input, and consumption data show the level of living by measuring what people acquire. Both can be defended as approximations to utility. The “indirect” utility function expresses welfare in terms of resources (positively) and prices (negatively). In practice this usually means income or resources deflated by a price index: real consumption or income, not money consumption or income. Whether consumption or income is measured, measures of prices are needed whenever analysts wish to compare people who face different prices, which will be whenever they make comparisons over time or space.

Another consideration about whether to use income (including income from assets) or consumption is the time period over which living standards are to be measured. At one extreme is a lifetime living standard, measured either by average consumption over a person’s lifetime or by the person’s total lifetime resources; apart from any bequests, these two concepts are the same. The issue here is that some poverty is only temporary (for example, students are poor in the short term but not over their lifetimes, while the elderly may be poor but have not been poor throughout their lives) so short-term measures of inequality can overstate lifetime inequality. One influential theory of consumption and saving is the “life-cycle hypothesis,” which asserts that a person’s consumption at any age is proportional to his or her lifetime resources. If this is true, measuring consumption is not only useful in its own right but also provides an indication of lifetime resources. However, the evidence for this hypothesis is controversial to say the least; for many people, the promise of resources in the future will do little to pay bills today. Policymakers have to deal with current poverty regardless of the long-term prospects of the poor; saying “Don’t worry, they will be OK later” about poor children or “Don’t worry, they’ve had their turn” about the elderly are not acceptable responses.

If a lifetime is too long a reference period, a day, a week, and a month are all clearly too short. Arguments can be made in favor of using a season as a reference

period; there is a substantial literature on seasonal poverty (see, for example, Sahn 1989). However, there seems to be a general consensus that a year is a sensible reference period over which to judge people’s living standards, even if this is inevitably a compromise that is too long for some purposes and too short for others. There is also a good deal of empirical evidence that even people in poor agricultural societies and people without the ability to borrow much can smooth their incomes within a particular year and perhaps over a series of years, so that consumption will reflect living standards at least throughout one year and perhaps over a series of years. (For a review see Bhalla 1979 and 1980, Musgrove 1978 and 1979, Paxson 1992 and 1993, Wolpin 1982, and Deaton 1997, chapter 4.)

If a year is chosen as the standard for assessing living standards but the survey in question can only hope to measure flows over a shorter period, consumption data will yield a more accurate estimate of living standards than will income data. Most people do not receive income every day, and many do not receive income every season—or at least not an equal amount every season. So while consumption over a week, two weeks, or a month is likely to be a reasonable indicator of living standards over a year or over a few years, income will not be. If analysts are interested in measuring averages, income variation will not matter much if the survey itself is spread over a year, since some people’s zero incomes will balance out others’ high seasonal incomes. However, analysts are usually interested not only in means—LSMS surveys are rarely the instrument of choice for estimating mean income or consumption—but also in inequality and poverty, which are sensitive to the tails of the distribution, especially the lower tail. Gathering data on the previous month’s income will overestimate inequality in annual living standards and, provided the poverty line is below the mode of the distribution, will overstate the fraction of people below the line. Although there are also random irregularities and seasonal patterns in consumption, they are typically smaller than those in income, because consumption is less tied to seasonal and weather-related patterns in agriculture than is income. Even so, consumption measured over a reference period of less than a year is likely to overstate poverty and inequality. In addition, the overstatement may not be constant over time if seasonal patterns change with time, because one year is different from

another—or over the long run, because agriculture accounts for a shrinking share of household income as economies become richer.

These arguments provide a persuasive case that, given the choice, (perfectly measured) consumption is a more useful and accurate measure of living standards than is (perfectly measured) income. These theoretical advantages of consumption are likely to decrease as the period over which it is feasible to gather data gets longer. If it is feasible to visit households on many occasions throughout the year this will clearly capture any seasonality in the household's income. Moreover, if the survey has a panel element so that income can be averaged over a series of years, it makes little difference whether income or consumption is measured, if one can be measured as accurately and as cheaply as the other.

**PRACTICAL ISSUES.** The choice between income and consumption is often determined more by practical considerations than by theoretical considerations. In the United States poverty is assessed by income, not by consumption; consumption cannot be used because the United States does not have a consumption survey of adequate size and quality to permit the estimation of the poverty numbers. In general, however, whenever a new or reformed survey is being planned, the designers have to choose whether to collect data on household income or consumption. Much hinges on the relative costs and relative precision of the data collection required.

Neither consumption data nor income data are easy to collect. For consumption, the need is for data on total household expenditure on goods and services. As will be discussed in the next section, these usually have to be gathered item by item. In some cases a substantial fraction of consumption does not come through the market, so imputations have to be made. In industrialized countries such as the United States and Britain, the detail and the associated time and effort of asking dozens or sometimes hundreds of questions often make it seem relatively more attractive to collect income data, especially in situations where income comes from one or two sources (for example, wages and pensions) that are easily recalled or for which independent documentation exists. By contrast, consumer expenditure surveys are seen as among the most “difficult and expensive surveys” to field in the statistical system (McWhinney and Champion 1974).

In the United States, the Consumer Expenditure Survey costs about five times as much per household as the Current Population Survey, which is the main source for data on income, earnings, and employment.

Even so, the concept of expenditure—giving money in exchange for a good or service—is clear both to interviewers and interviewees, whereas the concept of income, especially income from self-employment or own-business activity, is not. For own-account workers in agriculture and small businesses, their personal and business accounts are often hopelessly entangled. Thus, in agriculture and elsewhere, the only practical way to estimate income is to gather data on all transactions—business as well as personal—and to impose an accounting framework on the resulting information. This process is extraordinarily time-consuming, and the results are subject to large margins of error. Such difficulties in calculating income are not specific to developing countries; even in the United States and Britain, the various surveys—the Current Population Survey and the Consumer Expenditure Survey in the United States and the Family Expenditure Survey in Britain—do relatively poorly in gathering data on income from self-employment. (See Coder 1991 for the Current Population Survey, Branch 1994 for a comparison of the two U.S. surveys, and Atkinson and Micklewright 1983 for the Family Expenditure Survey.) The difference between developing and developed countries is that formal sector wages and salaries are much less common in the developing countries.

The income of many households—particularly but not exclusively agricultural households—varies seasonally throughout the year. In these circumstances, measuring households' annual income (which is the minimum amount of data needed to adequately determine poverty and distribution) would require many visits to the household or reliance on the ability of household respondents to remember their income from many months earlier. However, if consumption is smoothed over the seasons—and much of the literature already cited suggests that this is done in most households—consumption will vary less by season than income does. It may also be possible to collect useful data on annual consumption without making multiple visits.

It is generally thought that respondents are more reluctant to share information about their income and (to an even greater degree) their assets than about their consumption. Thus they are more likely to lie about

their income than about their consumption. In many countries income is taxable, at least in principle, and it may be hard for the survey interviewers to persuade respondents that the information they give will not be passed on to tax authorities. Rich households may refuse to grant interviews to the survey team and, if a rich family does grant an interview, the respondent, who may be a family member or a servant, will frequently be more knowledgeable about the household's consumption than about its sources and levels of income. Income from assets is likely to be particularly hard to capture because the ownership of assets is highly unequal, and the wealthy—who own the most assets—are typically thought to be the least likely to cooperate. Given that most of the survey interviews in developing countries must be conducted in a semi-public place, respondents are often reluctant to state their wealth in the presence of relatives and friends. These problems of measuring assets and asset income are likely more severe for measuring inequality than for measuring poverty, since households below the poverty line typically have few assets.

#### *What Analyses are Consumption Data Good For?*

The consumption data that can be gathered in LSMS surveys have a number of important analytical uses.

**MEASURING WELFARE.** The policy importance of measuring living standards is indisputable. Household budget analysis has been used to document and to publicize poverty since the late 18th century. While consumption generally cannot measure noneconomic components of living standards—health, access to education, political freedom—it is the best measure of the economic component of living standards. Formally, consumption is valuable as an approximation to utility, or “money-metric” utility, according to which an indifference curve is labeled by the amount of money at constant prices that is required to reach it (see Chapter 5 of Deaton and Muellbauer 1980). Total household expenditure adjusted by a price index and divided by the number of people in the household (or by some more sophisticated count such as the number of equivalent adults) is a measure of the living standard of each member of the household and is the measure recommended in this book for analyzing poverty and inequality. (See Deaton and Zaidi 1999 for a more comprehensive discussion.)

LSMS surveys also collect data on a wide range of other household and community variables that help

describe other dimensions of living standards. For example, they collect data on health outcomes and facilities and on educational attainments and facilities. These measures are frequently used not only to document living standards but also to explore their determinants in studies, for example, of the relationship between income, assets, and consumption, between earnings and schooling, or between health status, income, and consumption.

In addition to being used to construct a single summary measure of the economic welfare of households, the consumption data that can be collected in LSMS surveys have other important uses, some of which are discussed briefly below. For a much longer account with applications see Deaton (1997).

#### **EVALUATING THE IMPACT OF PRICE, SUBSIDY, AND TAXATION POLICIES AND THE PROVISION OF PUBLIC GOODS.**

Analysts are often concerned with the effects of price changes caused by changes in tax or subsidy policies or by fluctuations in world prices. Consumption data are invaluable for assessing these effects—in particular, who gets hurt by a price increase and to what degree. Many developing country governments collect a large share of their revenue through tariffs or through taxes on consumption, while simultaneously subsidizing the provision of many goods and services ranging from basic foods (such as bread, wheat, or rice) to transportation, health, and education. To a first approximation, a price increase hurts consumers in proportion to the amount of the good that they purchase, so in order to know the distributional effects of a price change, analysts need to know who consumes what and where consumers are in the overall welfare distribution. For example, do transport subsidies benefit the poor as is often claimed, or do they actually benefit people who are much better off? Improving the quality of clinics or increasing the number of teachers in schools will not help the poor if the poor do not use these clinics or attend the schools where these teachers are employed. Even simple cross-tabulations can establish results that, if not necessarily surprising, can resolve major policy controversies. (See Grosh 1997 on kerosene pricing in Ghana and health care use in Guyana; see Deaton 1988 on rice pricing in Thailand.)

More complex modeling of price reform requires estimates of how consumers respond to price changes,

so analysts can calculate dead-weight loss and the tradeoffs made between equity and efficiency. Once again, data on consumption, income, and prices that are needed to estimate these responses (Newbery and Stern 1987; Ahmad and Stern 1991; Deaton 1997, chapter 5).

**NUTRITION AND POVERTY LINES.** There is a long tradition in development economics of counting calories and of defining poverty in terms of malnutrition, for example by counting people whose caloric intake falls below some recommended standard. This tradition is misguided; nutrition is not an accurate measure of welfare because people consume more items than food and people often make tradeoffs between food and other goods. Thus collecting data on calories consumed is no substitute for estimating consumption. Nevertheless, documenting nutrition is of considerable interest for other reasons.

In some surveys a household's caloric consumption is estimated directly by nutritionists who enter the household and observe what is eaten by each household member, either by weighing and measuring foods as they are consumed or by asking the members questions about their dietary intake during the previous 24 hours. It is possible to imagine a module of this sort that could be added to a multitopic survey, although, because it would be lengthy, it might displace other modules in the questionnaire as a whole. Swindale (forthcoming) provides guidelines on how to collect dietary intake data.

Many writers believe that dietary intake surveys are necessary to obtain accurate estimates of caloric intake (Bouis 1994; Bouis and Haddad 1992). However, dietary surveys also involve a number of difficulties. The survey techniques are invasive and may cause people to alter their behavior. A household's day-to-day consumption may vary enough to make a 24-hour recall period too short to yield accurate data with which to estimate poverty, yet longer periods may be too expensive or too invasive. A more common—albeit probably less accurate—way to count calories is the “indirect” method, which is most often used in expenditure surveys. Figures for the quantity of each good that the household has consumed can be obtained in two ways. The survey interviewers can ask direct questions about both the physical quantity consumed by the household and the household's expenditure on the good, or they can collect data only on expenditures and

deflate these figures by the prices of the commodities in question (as obtained in the community or price questionnaires). Standard conversion tables are then used to convert quantities into a count of the number of calories contained in the food purchased—a measure known as “caloric availability.”

Data on caloric availability have been used together with data on household income or expenditure to calculate Engel curves that plot the average household caloric consumption at each level of income or expenditure. Following work that was done in India over 25 years ago (see Dandekar and Rath 1971a and 1971b and Government of India 1993 for a review), income or total expenditure poverty lines are obtained by calculating the income or total expenditure level at which the caloric Engel curve gives the recommended caloric intake. If the caloric Engel curve has a relatively high slope, increasing household income will eliminate hunger relatively rapidly. If, as some recent writers have suggested, the elasticity of caloric consumption with respect to income is close to zero, economic growth alone will not eliminate hunger. This means that poverty can only be reduced by direct intervention, an approach which is closer to the basic needs philosophy. (See Behrman and Deolalikar 1987 and Bouis and Haddad 1992—who also argue that estimates are biased when caloric availability is used rather than direct dietary surveys—as well as a contrary position from Subramanian and Deaton 1996. A review is provided by Strauss and Thomas 1995.)

The demand analysis discussed in the previous subsection can also be applied to calories to calculate the effect of price changes (caused by, say, the elimination of subsidies on basic foods) on caloric intake (Laraki 1989).

**INTRAHOUSEHOLD ALLOCATION AND GENDER BIAS.** Expenditure data are an important tool for researching the allocation of resources within the household and for testing different models of how that allocation might work. In recent years many studies have found different outcomes for males and females, particularly boys and girls, within the same household. In some countries infant mortality is higher among girls than among boys, and in even more countries educational outcomes are worse for girls than for boys. Several scholars have explored the possibility of using data on household expenditures to cast light on these different

outcomes for boys and girls, as well as to compare other pairings—adult women and adult men, the elderly and prime-aged adults, or widows and other household members. (For a fuller discussion see Chapter 24 on intrahousehold issues, as well as Deaton 1997, chapter 4.)

It is costly and time-consuming for surveys to collect complete data on every item consumed by every family member. In fact, this may be impossible for the many joint (or household public) goods that are shared by all household members. As a result, most multipurpose surveys, including the LSMS surveys, have collected household-level data on consumption and have made little effort to collect individual data. Nevertheless, there are some cases where consumption at the individual level can be inferred from household data. Such cases include health expenditures that are linked to an identified episode of illness on the part of one member or expenditures on men's clothing when there is only one man in the household. In some surveys data on expenditures have been collected using the diary method, in which each adult family member has been asked to keep a diary about his or her own expenditures, from which we can see who spends what in the household, even if not who consumes what. Even when data are collected by interviewers, it is probably possible to collect more individual data than has typically been collected in the past if the interviewer can find out who consumes how much of such obviously private goods as tobacco, transportation, clothing, or entertainment.

Even when individual-specific data are not collected, it is possible to examine the effect of household characteristics—including household composition—on the way households allocate their budgets. For example, it may be that household expenditures on food and children's clothing are higher when there are relatively more women in a household or when a large share of household resources are earned by, and thus putatively controlled by, women. There is also a developing literature (Bourguignon and Chiappori 1992; Bourguignon and others 1993; Browning and others 1994) that has identified sharing rules within the household. If some goods can be identified that are consumed exclusively by one group within the household or if analysts have data on who consumed how much of each good, it is possible to infer whether or not income is shared equally across the groups. Related to this is the examination of expenditures on

“adult goods” (usually alcohol, tobacco, and adult clothing) for signs of gender bias in the treatment of children. Since the total household budget is not increased by the presence of children, parents typically reduce their expenditures on adult goods to make room for the costs of the children. If the parents cut back on their own consumption more for their sons than for their daughters, this is evidence of discrimination against the girls in the household. Surprisingly, analysts have consistently failed to find such differences, even in places where there is other evidence of bias against girls—such as differential infant mortality (Deaton 1997, chapter 4).

**FAMILY STRUCTURE, CHILD COSTS, AND ECONOMIES OF SCALE.** The most commonly used measure of living standards is household total expenditure per capita—total household expenditure divided by the number of household members. This measure, while convenient, ignores the fact that the needs of one household member differ from those of another household member, particularly between adults and children, and that there are likely to be some economies of scale in household size. Larger households, which usually include many children, are likely to benefit most from economies of scale. Thus measuring living standards by per capita total household expenditure almost certainly overstates the number of large households that are poor and understates the number of small households that are poor. In some countries—most notably the United States—there is a different (official) poverty line for each type of household; these lines embody both economies of scale and the different needs of adults and children.

There is a long history of economics studies that have attempted to use consumption data to derive the cost of living for families of different types by inferring equivalence scales across age groups and estimating the extent of economies of scale. If such calculations were feasible and credible, they would have a key advantage over dividing resources by the number of people in the household because they would take into account country-specific and local differences in the costs faced by different types of families. For example, it is often argued that children are relatively more expensive in rich countries than in poor agricultural societies. Unfortunately, all procedures for estimating equivalence scales are controversial, and many economists would argue that the task is misguided or even

impossible. (See chapter 4 of Deaton 1997 for a discussion of both sides of the argument.)

Nevertheless, consumption data have a more limited but less controversial role to play in helping analysts check the implications of various models. Although all methods for measuring economies of scale or estimating equivalence scales must contain untestable identifying assumptions, most have stronger implications that can be tested using the data. The results can reveal a great deal about the plausibility of the models. For example, Deaton and Paxson 1998 used a number of LSMS data sets to show that the relationship between food expenditures and household size contradicts most of the obvious notions about how economies of scale might operate.

Without consumption data it is impossible to make any progress on the extremely important policy issue of how to factor differences in household size or structure into the assessment of household welfare. Until some agreed basis is established to correct for cost of living differences faced by households of different size and composition, there is no way to address such issues as the relationship between poverty and fertility or whether children are more likely to be poor than adults or the elderly.

**CREDIT AND SAVING.** A traditional use of expenditure data in analysis is to combine these data with income data to derive estimates of saving at the household level. The role that saving plays in economic development has always been an important intellectual issue, and both public and private saving are rarely absent from the policy debate. Unfortunately, the generally poor quality of data on savings collected through household surveys has limited their contribution to this debate, except perhaps in countries like Taiwan, where household saving rates are very high. Microeconomic income data have typically been poorly measured, and even if consumption measures tend to be more accurate, the estimate of saving is the relatively small difference between two large and inaccurately measured numbers, and as such, may be mostly measurement error. It is not clear that having such measures of saving is worth the effort of obtaining them. To the extent that it is the owners of small-scale, household-based activities who are doing the saving, it is even more difficult to measure saving because data on income from these activities are extremely hard to measure accurately.

Using data from an expenditure module to contribute to a measure of credit use may not be such a daunting prospect. Supplier credit constitutes a large share of households' total use of credit. A convenient way of eliciting information about supplier credit is to add questions in the consumption module about purchases on credit (see Chapter 21).

#### *Do LSMS Surveys Collect Accurate Consumption Data?*

In most developing countries there are no independent estimates of poverty and inequality against which LSMS data can be checked. However, it is possible to compare estimates of per capita consumption from the surveys with similar estimates from the National Income and Product Accounts (NIPA). Although the main purpose of the LSMS is not to measure means, if the LSMS data are enormously different from NIPA estimates, public confidence in the survey is likely to erode—particularly public confidence in the survey's estimates of consumption. While it is important to make these comparisons, it should not automatically be assumed that the NIPA estimates are correct and that discrepancies are wholly due to errors in the survey data. The quality of NIPA accounts varies widely across the world and, while some items of consumption are well estimated (for example, when consumption is from imports and there is good recordkeeping at the border), the data on other items are often no more than educated guesses (Srinivasan 1994). Even when this is not the case, there are frequently important differences in the definition of consumption between the NIPA and the household survey. If these definitional differences are not corrected for, the comparisons may not be valid. (For evidence from the United States see, among others, Gieseman 1987 and Branch 1994.)

Table 5.1 presents a number of LSMS survey estimates alongside their NIPA equivalents. (As far as the authors of this chapter are aware, LSMS survey data were not used to construct any of these national accounts.) This comparison should not be taken too seriously, for two reasons. First, no detailed investigation of NIPA practices for the countries in the table has been undertaken, so there is no rigorous information about the accuracy of their estimates. Second, the survey numbers were taken from the various survey reports rather than from the original microeconomic data (which would have been prohibitively expensive). As a result, there may be some incomparabilities in their calculation.

**Table 5.1 LSMS and NIPA Estimates of Average Per Capita Consumption, Selected Surveys**

Country	Dates	Currency	LSMS annual mean per capita expenditure	NIPA annual per capita consumption	Ratio of LSMS/NIPA	Sources
Bulgaria	5/95–7/97	levas	50,436	90,021	0.56	Authors' calculations from data on LSMS Web-site; staff estimates
Côte d'Ivoire	2/85–1/86	CFA francs	237,853	184,935	1.29	Grootaert 1993, p. 30; IMF 1995
Côte d'Ivoire	2/86–1/87	CFA francs	223,905	194,554	1.15	Grootaert 1993, p. 30; IMF 1995
Côte d'Ivoire	3/87–2/88	CFA francs	216,965	190,032	1.14	Grootaert 1993, p. 30; IMF 1995
Côte d'Ivoire	5/88–4/89	CFA francs	173,072	190,203	0.91	Grootaert 1993, p. 30; IMF 1995
Ecuador	6/94–9/94	suces	2,032,560	2,230,392	0.91	Lanjouw and Lanjouw 1996, table 3; IMF 1996
Ghana	9/87–8/88	cedi	56,645	45,568	1.24	Glewwe and Twum-Baah 1991, p. 17; IMF 1995
Guyana	1/93–11/93	Guyanans dollars	91,602	53,750	1.70	World Bank 1994a, p. 7; Baker 1996
Jamaica	8/88–9/88	Jamaican dollars	4,700	5,210	0.90	World Bank 1996, p. 28
Jamaica	11/89–3/90	Jamaican dollars	6,304	6,568	0.96	World Bank 1996, p. 28
Jamaica	11/90–4/91	Jamaican dollars	7,616	7,869	0.97	World Bank 1996, p. 28
Jamaica	11/91–2/92	Jamaican dollars	10,384	11,092	0.94	World Bank 1996, p. 28
Jamaica	8/92–3/93	Jamaican dollars	16,998	17,718	0.96	World Bank 1996, p. 28
Jamaica	11/93–3/94	Jamaican dollars	23,408	23,684	0.99	World Bank 1996, p. 28
Jamaica	11/94–1/95	Jamaican dollars	32,712	35,819	0.91	World Bank 1996, p. 28
Kyrgyz	10/93–11/93	som	2,273	907	2.50	World Bank 1995a, p. 60
Morocco	10/90–11/91	dirham	6,870	6,384	1.08	World Bank 1994b, volume II, annex 1, table 2
Nicaragua	2/93–6/93	cordobas	4,079	2,312	1.76	World Bank 1995b, volume II, p. 46; IMF 1996
Pakistan	1/91–12/91	rupees	6,835	6,037	1.13	Lanjouw and Lanjouw 1996, table 4; IMF 1995
Peru	7/85–7/86	intis	4,616	6,359	0.73	Glewwe 1987, p. 9; IMF 1995
Peru	10/91–11/91	new soles	750	1,178	0.64	Webb and Baca 1993, p. 266; IMF 1995
Peru	7/94–8/94	new soles	2,190	3,539	0.62	Authors' calculations from data on LSMS website, 1996; IMF 1996
Romania	4/93–12/94	lei	1,126,558	1,348,055	0.84	World Bank 1987, annex 1, table 4; National Commission for Statistics
Russia	10/93–2/94	rubles	1,071,312	497,512	2.15	Foley 1996; IMF 1996
Tanzania	9/92–11/93	schillings	129,708	37,718	3.44	World Bank 1995c, p. 30
Venezuela	1/92–12/92	bolivares	69,684	142,104	0.49	Scott 1994, p. 15; IMF 1995

Note: Adjustments were made for inflation as follows. If the country had less than 15 percent inflation, and the survey period covered the whole year, no adjustment was made. If the country had less than 15 percent inflation and the survey period covered only part of the year, the survey figures were adjusted to correspond to midyear prices. If the country had greater than 15 percent inflation, the monthly consumer price index was used to adjust the NIPA numbers to the month for which the survey's data are priced.

Source: Author's summary

The table shows average per capita consumption for 26 LSMS surveys. The ratio of the LSMS to the NIPA estimate has a median of .96. Though these summary measures indicate an impressive consistency between the survey and NIPA estimates, there are large discrepancies for some countries and years. Nevertheless, the survey estimates provide no evidence that expenditures are generally understated—defying the common belief among survey experts (backed up by substantial literature) that the understatement of expenditures is the major problem with

consumer expenditure surveys in industrialized countries.

Looking at some of the cases with large discrepancies between the household survey data and the NIPA estimates illustrates the difficulties involved in making comparisons between these surveys. Several of the countries surveyed were undergoing major economic changes at the time. For example, the som had been introduced as a national currency just five months before the Kyrgyz survey was carried out in October 1993. While the ruble had lost its status as

legal tender in the Kyrgyz Republic, it was still used in the Republic's neighboring countries, with which it had substantial formal and informal trading relationships. In this rather chaotic situation, inflation rose to 772 percent per year, and the dollar became a *de facto* unit of account—the most reliable store of value and sometimes the unit of transaction for large purchases or purchases of imports. The planners of the Kyrgyz survey had been perplexed about whether to use the som, the ruble, or the dollar as the survey's unit of account and about what would constitute reasonable recall periods, especially for nonfood items.

The same changes that complicated surveying also complicated national accounting in the Kyrgyz Republic. The national statistical office had just begun to calculate the international NIPA numbers rather than the Gross Social Product numbers that used to be calculated in the Soviet Union. With such complications in the measurement of both numbers, it is not surprising, nor even particularly alarming, that the survey data are not very close to the NIPA calculations for the Kyrgyz Republic. Similar (though somewhat less dramatic) issues plague the comparisons for Russia, Romania, and Bulgaria, also contributing to very high discrepancies between LSMS and NIPA estimates.

An anecdote about Guyana further illustrates the problems of comparing household survey data with NIPA estimates. When the first tabulations were becoming available from the survey, they yielded a mean per capita expenditure about twice the amount estimated in the official NIPA figures. This caused some consternation among the team working on the survey analysis. However, within hours, they discovered that an effort was concurrently being made to adjust the NIPA for various flaws and biases. The new estimate for NIPA was within \$30 per capita of the survey data (van der Gaag 1994).

The teams working on the surveys in both Tanzania and Nicaragua also noticed that the estimates of consumption from their surveys were much higher than the estimates in the national accounts. However, they were not disturbed by this, because national accounts in those countries had a reputation for being inaccurate (Tsoflias 1996; Scott 1997).

As can be seen in Table 5.1, LSMS estimates for Venezuela were lower than NIPA estimates. In this case, it appears to be the survey that was wrong. A very short list of consumption items was used in the

Venezuelan survey, and, as a result, consumption was probably underestimated (see the next section). Why the Peruvian surveys' consumption estimates were substantially lower than those of NIPA is a puzzle, although it may have been that the recall periods used were too long given prevailing rates of inflation. Yet the sample for the 1991 LSMS survey in Peru omitted some sizable rural areas, which means that the survey estimate should have been higher, not lower, than the NIPA estimate.

### **The Data Requirements of a Consumption Survey and How to Meet Them**

This section discusses the data needed to obtain a consumption-based measure of living standards as well as to analyze the other policy and research issues outlined in Part 1. To the extent that this section focuses on the main purpose of this chapter—the measurement of a single aggregate for consumption—the discussion is driven less by what to measure than by how to measure it. In this respect, this chapter differs from many of the other chapters in the book. This section discusses conceptual issues, the measurement of prices, and the design issues that have loomed large in analyses of LSMS data and in the previous literature. Thereafter, the section presents the current best practice on each of the issues, based in part on the consumption literature but also on LSMS experience—and results from experimental surveys that were specifically conducted to study points discussed here.

This task has been a difficult one, and there are several important issues on which there is little to report. Even for consumer expenditure surveys in industrialized countries, which have been extensively documented and which have featured a good deal of experimentation, the literature has not produced a satisfactory synthesis between theory and practice. The literature is difficult to find and is scattered across various disciplines, including economics, marketing, psychology, sociology, and statistics. Much of it is contained in poorly catalogued government reports and conference proceedings, rather than in academic journals. Even so, there has been a good deal of recent progress in understanding how to measure consumption, particularly from interactions between survey statisticians and cognitive psychologists. There are now several topics, reasonably well understood by analysts, in which conclusions can be extrapolated from previ-

ous experience with some confidence (see Sudman and others 1996 and the 1991 volume edited by Biemer and others). The discussion presented here will draw on this work.

Even so, it should be kept in mind that there are many design issues where there is evidence of problems but little understanding of their causes nor solid recommendations for solutions. As Sudman and others (1996) point out, “The theoretical basis of interviewing is still less rigorously developed than the theoretical basis of sampling.” Thus the wording and design of questionnaires has largely remained an art governed by in-house tradition and personal experience. In his useful discussion of the sources of measurement error in expenditure surveys, Neter (1970) gives the following typology of (nonsampling) errors that remains relevant today (with applicability that goes well beyond consumption data):

- Recall errors associated with the fading of people’s memories.
- The “telescoping” of reported events by incorrect dating.
- Reporting errors associated with respondents being overwhelmed either by the length of the survey or by the number of items covered.
- “Prestige” errors—in other words, misreporting due to various social pressures.
- Conditioning effects from being in the survey.
- Respondent effects, in which the respondent’s identity affects the answers he or she gives.
- Interviewer effects.
- Effects associated with the design of the instrument.

Another nonsampling error that could be added to Neter’s list is biases in the data due to nonresponses or to the use of an inadequate sampling frame. All of these issues have been extensively discussed in the subsequent literature, but only for the first two—fading memory and telescoping—are the causes and treatments reasonably well understood 30 years later.

The literature on expenditure surveys in developing countries is even thinner. While findings from rich countries often carry over to poor countries, there is always reason to be cautious. For example, response rates in LSMS surveys are very high—nearly always higher than 80 percent, and often closer to 100 percent. This is higher than response rates for surveys in many developed countries; response rates are 85 percent for the U.S. Consumer Expenditure Survey, about 70 percent for the British Family Expenditure Survey,

and typically lower for surveys in much of western Europe. As a result, if the difficulty of surveying wealthy households is one factor that causes average consumption to be understated in industrialized countries, there is likely to be less underestimation from surveys in poor countries. While the high response rates are good in themselves, those who choose to respond will not necessarily be attentive and cooperative, say, in terms of tolerating long questionnaires or taking enough time to try to remember the information they are asked to provide.

The LSMS program of surveys has contributed little to the methodology of measuring consumption. Systematic experiments have been conducted only recently and are still largely not written up in places easily available to the community of survey researchers. The emphasis on collecting data rather than on furthering methodology may be inevitable given that the countries and country departments (within the World Bank) that fund the surveys are more interested in increasing their understanding of development policies than in increasing survey know-how. Nonetheless, it is lamentable that the LSMS program has not done more in this area and desirable that it should do more. Good survey practice should include the continual evaluation of methods, just as good social policymaking includes evaluation of the impact of government programs. The LSMS program has been remiss by not more strongly supporting investigations into survey methodology.

Even while maintaining its principal focus on data production, the LSMS could make useful contributions to survey methodology in four areas. First, it would be helpful and not too costly for the LSMS survey teams to produce fuller documentation of their pretests of questionnaires, including the different options that were tested, the process and lessons of the field test, and the reasons for making the choices that were made about the questionnaire. A great deal is learned in this process and never made systematically available to others. The record might include taping of interviews and debriefing of respondents about the interviews.

Second, by more rigorously collecting and disseminating metadata on the process of interviewing (as described in Chapter 4), analysts will be able to systematically study survey methods, costs, and quality. Third, more controlled experiments with alternative modules—such as those that were set up or analyzed

as part of the background research for this chapter—should be done. These are particularly appropriate in countries where the new survey will use a different consumption module than has been used in past surveys. In such cases the experiments would not only increase knowledge about measurement of consumption but would also allow the host country to make adjustments in comparisons between its older surveys and the new one, so that the differences in results due to method could be identified.

Fourth, although the consumption modules used in most surveys are increasingly standard, there is much to be learned from talking through design issues in “cognitive laboratories” in survey organizations (see Sudman and others 1996). In these sessions, potential respondents are brought into a laboratory, asked sample questions, and then de-briefed on how they interpreted the questions and how they went about answering them. The results of these sessions are then used to modify the questions in the module as necessary. Sudman, Bradburn, and Schwarz (1996) wrote, “In questionnaire design, we strongly recommend the use of think-aloud interviews for determining what respondents think the questions mean and how they retrieve information to form a judgment.”

#### *What Consumption Data are Needed?*

The measure of living standards that analysts wish to construct is a real value of total household consumption on a per capita or per equivalent basis. Thus they need to have available information on three things: consumption, household size (including—for the equivalence scales—the age and sex of household members), and prices. Data on household size (and age and sex of household members) can be gathered in the household roster (see Chapter 6); consumption and prices are discussed here.

A measure of total household consumption is built up from several components. First it is necessary to add up all reported expenditures on individual goods and services or groups of goods and services (for a fuller discussion see Deaton and Zaidi 1999). Then a value for consumption that does not go through the market—in other words, consumption out of home production or in-kind received from employers—must be added in. In countries where households hold significant stocks of goods, particularly expensive durable goods, it is necessary to correct for the difference between consumption and expenditures. Then

the estimates must be converted to real terms by adjusting them by a price index to account for differences in prices among different regions or interview dates. It is important to note that this accounting should be done after the data are collected because it is neither necessary nor advisable for the respondent to understand the economic concept of consumption (or of income). The questionnaire should be designed around items familiar to the respondent—typically, cash flows or flows of goods—while gathering enough information to allow total consumption to be calculated. However, consumption questionnaires often collect additional data on cash flows that are not part of the economist’s definition of consumption but that are of interest in themselves and that the respondent sees as outlays of cash similar to purchases of goods. Such cash flows include taxes, contributions to savings accounts, or loan repayments.

To measure welfare accurately, the consumption concept must be comprehensive. All goods and services that contribute to people’s standard of living need to be included in the measure, which can be thought of as a practical approximation to an indirect utility function or money-metric measure of welfare. While it is often tempting—and economical—to collect data on only a subset of consumption (or sometimes even a single good or group of goods, such as housing or food), the relationship between the part and the whole can vary a great deal from one household to another and from one place or time to another, so that rankings or living standards obtained using the shortcut measure may not be universally valid. A good example comes from the spatial differences in relative prices that cause people to substitute cheaper goods for relatively more expensive goods. Poor urban dwellers must often live in poor housing in order to have access to income-earning opportunities in the city, but the standard of their housing will understate their overall standard of living.

As will be seen in the following sections of this chapter, there is not a clearly “right” or “wrong” way to resolve many issues about how to measure consumption. Rather, there is a range of good practice techniques and not enough empirical evidence about which is best. However, it is true that the estimates of consumption are sensitive—sometimes markedly, sometimes only slightly—to which method is used to formulate the estimate. In addition, survey designers will often wish to ensure that the new survey data are

comparable with previous survey data in that country, which is a powerful argument in favor of whatever method was used in all (or most) of those previous surveys. However, this may sometimes conflict not only with the interests of accuracy and best practice but also with standards that would allow the survey data to be compared with equivalent data from other countries. Comparability over time within a country is useful for monitoring poverty, which often takes precedence over other considerations. In addition, international agencies and researchers value international comparability, and because they often provide the funds or technical assistance for surveys in developing countries, this consideration is frequently influential. Even so, the past survey practices in a given country may be far from sensible or standard; when this is the case, it may be better for a new survey to be the first in a new series of consistent and potentially comparable surveys than for it to replicate a flawed method of measuring consumption. At the very least, experiments should be conducted to test whether the previous surveys actually collected the data they purported to collect. Comparability of nonsense is no great virtue.

For measuring welfare, consumption is ultimately a more useful measure than expenditures (purchases). For most perishable goods it is safe to assume that a person's or household's consumption is closely tied to their purchases. A kilo of tortillas or a bunch of bananas must be consumed soon after they are purchased; for such goods, expenditure and consumption will approximate each other over a short period. Even for less perishable commodities, some averaging across goods may occur in a fairly short period of time. A person may buy a pound of coffee one week and consume it over a month, but the next week he or she may purchase a pound of sugar that will also last for a while, the next week he or she may purchase a bag of flour, and so on.

In the case of major durable goods, expenditures and consumption are not closely related in the short run and household expenditures on durable goods will be a poor guide to the consumption of durable goods. (In some cases, where grains can be stored for substantial periods of time, the same may be true for goods that are not conventionally classified as durable. Rice or dried pasta may be stored for a long period of time, so their consumption and the expenditures made to purchase them may deviate considerably.) For major

durable goods (and in some cases for stocks of grain or fuel), consumption should be linked to stocks rather than purchases; thus the submodule that deals with durable goods needs to compile a list of the household's durable goods. Some sort of consumption flow needs to be imputed from this list. To impute the consumption flow sensibly, analysts need to know both the age of the good and its original (and perhaps current) value. In the case of housing that has no adequate rental market, analysts need to know any characteristics of this good that can be used to impute its rental value. Of course, such imputation is at best a hazardous undertaking in countries where there are few rental units to judge by, and the quality of the resulting data may not be worth the effort to collect them. Great care must also be taken to avoid erroneous interpretations of the results in cases where such imputations have an important effect on the total consumption measure or on the welfare rankings of households. If there is no rental market and possibly only a limited housing market, an imputed rental value may overstate the value of the housing to its inhabitants. Particularly in an emergency, it may be hard if not impossible for them to turn this imputed value into urgently needed cash. It is unwise to let policy decisions rest on often arbitrary and contentious imputations.

The policy issues discussed in the first section can mostly be analyzed with data that are the byproduct of the need to construct an estimate of total household expenditure. There are, however, some exceptions where more data will need to be collected to allow analysis of a particular issue.

First, the decision about the level of disaggregation at which to collect the data must be guided by the needs of analysts to have data on specific items of expenditure. For example, an analyst might wish to have data on goods of particular nutritional significance, such as rice or milk. Likewise, when different goods are taxed or subsidized at different rates and analysts want to use the survey data to investigate tax reform, the different goods must be distinguishable from each other. If flour is subsidized in the country where the survey is to be fielded, then it would be a good idea to include a separate question on the consumption of flour rather than including it in a broader question about "staples" or "flour, rice, and cornmeal." Where analysts are concerned with the relationship between consumption and the environment, it is necessary to distinguish consumption items

that were gathered or hunted from consumption items that were grown at home or bought in the market.

Second, analysts may wish to collect data that are disaggregated so as to yield information about intra-household allocations, even if such disaggregation is not required to estimate the total of all expenditures. Items that are exclusively consumed by different groups are an obvious example; for example, rather than having a single item for clothing, data on men's and women's clothing and footwear can be collected separately from data on children's clothing. For at least some goods it is also possible to include questions about who consumed what. Tobacco and alcohol are consumed individually, not jointly, and it is probably possible to obtain reasonable estimates of the individual-level consumption of these products. Other examples of goods that are consumed individually are tickets for entertainment and transportation. One difficulty is that these may be the items for which the best-informed overall respondent for the household gives the least accurate answers. For example, the homemaker may know a lot about the household's food consumption but relatively little about individual members' expenditures on alcohol, tobacco, or entertainment. This problem can be dealt with by conducting individual interviews on consumption of some goods—a feasible but costly solution.

Third, some of the research topics, such as the calculation of calorie availability or the estimation of price elasticities, ideally require data on quantities of individual items at the household level, whereas the welfare measure makes do with expenditures, deflating when necessary by a price index. The issue of quantities will be addressed in the next subsection.

#### *Collecting Price Data*

The conversion of money values to real expenditures requires the construction of a price index; to construct a price index, price information must be available. This price information must not only capture temporal variations in prices but also accurately represent the price level faced by each of the households in the survey sample. While adequate price indices may already be available in some countries, such cases are few and far between because many price surveys exclude rural areas. Urban prices are only useful for nationwide analysis when spatial price variation is limited—for example, where there is a good transportation network and markets are well integrated (although even in this

case there may be marked regional variation in housing costs).

When there is no other adequate price information, data must be collected in the household survey. This may be done at the household or community level. At the household level, the survey may be designed to ask each household how much was paid for each unit of an item purchased, the quantity of the good purchased, and total household expenditures on the good. When households report physical quantities (such as kilos, sacks, or numbers), it is possible to divide their reported expenditure by the reported quantity to yield a price—or more precisely a unit value—for each good; these values can be weighted together to create household-specific price indexes.

In most previous LSMS surveys, designers opted to collect price and quantity data from local markets in a community-level questionnaire, with few LSMS surveys collecting such information from households. However, many other surveys around the world collect quantity information from households—including the Indian National Sample Survey, the Pakistan Household Income and Expenditure Surveys, and the Indonesian National Socio-Economic Survey (SUSENAS). While the LSMS survey in Vietnam did collect price data from local markets, it also collected quantity data at the household level. This was also the case for LSMS surveys in Brazil, Ecuador, the Krygyz Republic, Nicaragua, and Russia. The LSMS surveys in Pakistan, Bulgaria, and Ecuador included questions about expenditures per unit. Apart from the Pakistan survey, where the unit cost data had serious problems (probably for local reasons, almost certainly from inadequate interviewer training), so far there has been no systematic evaluation of household-level price and quantity data collection.

There are several advantages to collecting price data by asking household respondents about their expenditures and the quantities of their purchases. This procedure yields measures of physical quantities that are useful in their own right for such purposes as computing calorie availability or estimating the elasticity of quantities relative to changes in taxes or subsidies. It also yields the raw material for a price index for each household without requiring the formulation of assumptions about where the household buys its goods. A price index constructed in this way, however, covers only those goods—typically but not exclusively foods—for which quantity data can be well

defined in the questionnaire. Such price indexes are automatically tailored to the consumption patterns of the households in the survey, so there is no discrepancy between the price data and the goods that people buy. Having price data for individual goods at the household level is also useful for analyzing demand patterns and policy issues—such as price reform—that depend on the results of demand analysis. One of the authors of this chapter has satisfactorily matched household data on unit costs from the Indian National Sample Survey to the prices that the Government of India regularly collects from local markets around the country, at least in cases where the local markets are located near the survey households. Not only do these data match across districts, but the unit values from the survey reflect the appropriate seasonal patterns of agricultural prices (see Deaton 1997, chapter 5).

There are also disadvantages to collecting price data this way. Unit values are not prices, and they vary even among households that purchase from the same sources, because better-off households typically buy higher qualities even of fairly homogeneous commodities like rice or sorghum, and certainly so for heterogeneous categories such as meat. One way this problem can be dealt with is by averaging the unit values over all the households in a primary sampling unit; the Indian evidence quoted above suggests that averaged unit values are not likely to be misleading as indicators of price. A more serious problem is that, with a few exceptions such as fuels and tobacco, it is not easy to define physical units for goods other than foods. Gathering data on the price of food may be enough in some cases—particularly in very poor economies where food consists of two-thirds or three-quarters of the budget of most households—but this is clearly not true in general. In many past LSMS surveys, even the definition of physical units of some food items was unclear or subject to error. For example, the respondents may not have understood whether to report the price they paid per egg or the price they paid for a dozen eggs. Moreover, goods are often sold locally in amounts or units that are often not very precise and thus can be hard to interpret at the analytical stage; a “bunch of vegetables” is much less clearly defined than a kilogram of rice.

LSMS surveys have had less difficulty in defining useful units for consumption than for production, where respondents often cannot provide any precise quantities (saying, for example, “I sold three sheets of

this and five baskets of that”). Since most of these difficulties were experienced in Africa and since successful experiences cited regarding quantities are largely Asian (though they also include Latin American countries and countries of the former Soviet Union), there may be a “continent effect” here, perhaps reflecting the degree to which the economy is monetized or to which nonstandard prices or units prevail.

The alternative to collecting prices from households is to collect prices at the community level at village and local markets in the primary sampling unit. This option is cheaper because prices are collected only for each primary sampling unit and not for each household. This option also has the advantage that, in principle, the prices in the market are the prices that consumers actually face. The fact that observed prices are the same for everyone in the primary sampling unit is thus an advantage, not a disadvantage. Most countries have some sort of regular method for collecting data on consumer prices and aggregating them into a price index, at least in urban areas; survey designers may be able to make use of or at least adapt these well-established procedures.

However, there are also a number of difficulties with collecting price data at the community level. One problem is that in some circumstances it is difficult for a survey team to replicate the sort of transactions that locals engage in; haggling is often an important factor in defining the prices actually paid by local consumers, which may mean that the prices vendors quote to survey enumerators are different from those paid by long-standing or regular customers.

A second problem is that the price questionnaire can only collect prices on items that are available in the local markets—which may exclude many nonfood items as well as those food items only consumed seasonally, or consumed regionally rather than nationwide. To solve this problem as well as the problem of defining a suitable unit, price collection in the survey may be biased toward manufactured or processed items that can easily be defined, such as a can of standard-brand tomato paste, a two-gallon plastic bucket, or a two-pound packet of sugar from the national refinery. Geographic problems remain. In countries where consumption patterns differ radically across regions (for example, between northern and southern areas of India), there may be conceptual difficulties that are as serious as the difficulties involved in comparing prices across countries.

A third problem is that in some (but not all) countries, it can be hard to know what is meant by a “local market.” The image of a primary sampling unit as an isolated rural village with a single market is an appealing one, but it is not accurate in all parts of the world. In urban areas, people may buy goods far away from where they live. (The obvious solution to this problem would be to use the urban price indices that are often available and that are sometimes of acceptable quality.) Even in rural areas primary sampling units are defined by statistical and often ultimately administrative criteria that may not accurately represent actual villages or village markets. So there is no guarantee that the prices in any given local market are the prices actually faced by households in the survey. This problem can be exacerbated when there are several different types of outlet in a given community—such as markets, corner stores, supermarkets, and subsidized government ration shops—all with different prices and clienteles.

Fourth, in at least a few past LSMS surveys, the procedures for entering data from the community questionnaire into the computer were not well enough established, and the data from many questionnaires or (in extreme cases) the whole community survey were lost to analysts. The recommendations in the new LSMS implementation manual by Grosh and Muñoz (1996) should overcome some of the problems, but community questionnaires will continue to be more novel to most survey agencies than household questionnaires, and therefore may be managed less well.

In the LSMS surveys for Pakistan (1991) and Vietnam (1992–93), price data were collected at both the household level (on expenditures and quantities in Vietnam and on quantities and unit values in Pakistan) and the community level, and the results for various goods were compared. The comparison is clouded by the fact that, for the reasons given above, the two data collection procedures measured different things, so that it is unclear exactly how close the unit values from the household questionnaires could be expected to be to the market prices from the community questionnaires. For both countries the two sets of estimates are similar at a sufficiently high level of aggregation; for example, there is little difference between the unit values and prices at the all-province level in Pakistan. However, the unit values and prices in both countries’ data sets differ markedly at the primary sampling unit level when one compares the estimate from the price

questionnaire with the average of households in the primary sampling unit. In Vietnam the correlations between the reported unit values from market purchases and the directly observed market price vary from 0.77 (for noodles) and 0.76 (for pork) to  $-0.07$  (for cassava) and  $-0.34$  (for mangoes). For the 16 foods in the comparison, cassava and mangoes are the only two negative correlations, and the median correlation is 0.34. The correlations are similar, if somewhat lower, for Pakistan—perhaps reflecting the problems with the unit value data.

Since any household survey requires some estimate of price variations, and since households cannot usually provide price (or unit value) data for most nonfoods or even for some foods, the survey always has to include a price (community) questionnaire of some kind. If the household questionnaire does not collect data on quantities, meaning that it will yield no unit values, the community questionnaire must gather data on the prices of food as well as nonfood items. When, as for many foods, both approaches are available, it is unclear from these two examples or from the literature whether the community or household method is preferable for any given survey. Of course, duplicating the collection of some data is an insurance policy; in at least one LSMS survey respondents were mistakenly given the option of reporting either quantities or expenditures, so that without the community questionnaire it would have been difficult to construct the essential consumption expenditure aggregate for each household.

It is useful for analysts to have data on quantities for reasons that go beyond the construction of price indices. Thus most surveys with objectives similar to those of an LSMS survey will want to collect data on quantities if it is feasible and economical to do so. If previous surveys in a country have yielded good data on quantities, or if it proves possible to gather data on quantities in an exploratory field test, questions about quantities should be included in the consumption module of the household questionnaires—at least for quantities of food.

#### *Level of Disaggregation*

The number of items about which data are collected is one of the central issues in designing the questionnaire for a consumption module. On the one hand, longer consumption modules are more costly and crowd information out of other modules of the ques-

tionnaire. On the other hand, asking about more specific items in detail is generally assumed to yield fuller reporting and greater accuracy than would asking about shorter, more general lists of items. There have, however, been suggestions that a survey can try to gather data in too much detail. Respondents may become bored or despondent. Or, believing that they are being uncooperative or showing themselves to be inadequate consumers if they have nothing to report in response to a long list of questions, respondents may invent purchases to be “helpful” or to enhance their prestige. However, there has been a validation using crop and trade data of food estimates from the Indian National Sample Survey which uses a very long list of items (Minhas 1988; Minhas and Kansal 1989), and, although there is some slight evidence of overestimation for cereals (about 3 percent), this seems well within reasonable bounds.

Traditional expenditure surveys in developing countries (for example, the National Sample Survey in India) use very long lists of consumption items, naming each food with great specificity. Lists of 200–300 items are not unknown, and the Brazilian budget survey uses a list of 1300 items. LSMS surveys, however, have been less detailed; the 33 foods and 20 nonfood items listed in the Pakistan survey in 1991 and the 45 foods and 46 nonfood items listed in the 1993–94 Vietnam survey are typical. As has already been discussed, some disaggregation is necessary to obtain information on certain specific items of interest; however, the precise level of disaggregation that survey designers choose for any given survey will depend on their views about the tradeoff between costs and accuracy.

There is a good deal of debate about whether short (or at least shorter than standard) consumption questionnaires can save time and money and still deliver accurate estimates of total consumption. The issue does not seem to have been settled in the literature. One set of results suggests that short lists of items will yield reasonably accurate data. A study by Reagan (1954) of farm operators in the United States found that total consumption was only modestly lower—about 10 percent overall—for a condensed list of 15 items than for a list of over 200 items. For developing countries, Bhattacharya (1963) reported on a small-scale experiment on 44 households in two villages in West Bengal that were presented with the usual detailed National Sample Survey questionnaire as well as with a questionnaire covering broad commodity

groups, along with a single question about their total expenditure during the previous month. The consumption totals from the questionnaire with the broad commodity groups were only slightly lower than the totals from the detailed questionnaire, while the single-question method gave consumption estimates that were 25 to 30 percent lower than the estimates from the full NSS list—although still highly correlated (0.98).

There is also some different evidence from the United States Consumer Expenditure Survey, in which certain households are asked to keep detailed product diaries of their purchases of food, while others (who are in the interview section of the survey for which food is not the main focus) are asked to report their total expenditures on food at home and away from home for each of the previous three months. According to Gieseman (1987), the amount that respondents said they spent on food at home was larger in the interview part of the survey than in the diary part of the survey—and also closer to the NIPA estimate. For the estimates of the amount spent on food away from home, the diary appears to be quite close to NIPA’s estimate, whereas this expenditure was substantially underestimated in the interview part of the survey.

Even more positive results were reported by a World Bank (1993) test survey in Indonesia. Both a short and a long questionnaire were administered to 8,000 households. In the short questionnaire the number of food items was reduced from 218 (in the long questionnaire) to 15 and the number of nonfood items was reduced from 102 (in the long questionnaire) to 8. While total measured food expenditures differed little between the questionnaires in terms of means and distribution, the long questionnaire showed about 15 percent more nonfood expenditure.

These results have not been replicated elsewhere. A similar experiment in El Salvador which reduced food items from 72 to 18 and nonfood items from 25 to 6 resulted in ratios (long-to-short) of 1.27 for food and 1.40 overall (Jolliffe and Scott 1995). A 1994 experiment in Jamaica, comparing modules with a total of 119 items to modules with a total of 37 items, produced a long-to-short ratio of 1.26 for both food and nonfood items (Statistical Institute and Planning Institute of Jamaica 1996, appendix 3). In Ecuador in 1993 two versions of the submodule on food items were piloted, one with 122 food items and the other

with 72. The ratio of total food expenditures reported in the long module to those reported in the shorter module was 1.67 (see Steele 1998). Shorter questionnaires sometimes dramatically reduce both survey costs and times compared to the longer questionnaires; in West Bengal survey time dropped from 180 minutes to 90 and in Indonesia it fell from 80 minutes to 10. However, it seems that such savings are often gained at the expense of accuracy.

There are alternatives to having either a long questionnaire or a short questionnaire. One alternative, which has never been used in an LSMS survey, would be a hierarchical scheme in which respondents are asked if they have purchased anything within a broad class of goods, and they are only asked the detailed questions about items in that class if they say they have bought a good in that class. For example, if a household responded that it had not bought any dairy items, it would not be asked specific questions about milk, yogurt, butter, and ice cream. Something like this may already happen in practice when interviewers and respondents are faced with lists of several hundred items that are inevitably grouped into broad categories. This kind of approach has obvious advantages when interviewers use computers to administer questionnaires to respondents, in which case the detailed questions about each item would never come up in the interview unless the respondent first indicated that they were relevant. The risk is that the categorical approach will cause consumption to be underestimated because respondents forget purchases that they might remember if they went through the list in detail. There is evidence from diary surveys in industrialized countries that preprinted diaries identifying more categories of consumption cause respondents to report more consumption, presumably due to this prompting effect (Tucker 1992; Tucker and Bennett 1988).

There are many different ways of turning a list of several hundred items into a list of only a few dozen. The traditional procedure, which might be referred to as the “botanical” method, groups together cereals, as well as pulses, root vegetables, or leafy greens. Since botanically similar foods often contain similar amounts of calories per kilogram, this way of aggregating the list of items ensures that analysts can calculate calorie counts when necessary. Other criteria can also be used for aggregating these items, such as where the consumer typically buys the goods. Consumers may more

accurately remember how much they spent in individual stores than how much they spent on individual goods or groups of goods. One experiment in Jamaica compared botanical aggregates and “point-of-purchase” aggregates and found little difference between the two sets of means and variances (Statistical Institute and Planning Institute of Jamaica 1996, appendix 3). There is also related evidence from the literature on diaries; Sudman and Ferber (1971) tested diaries that used itemization by purchase, groupings by product type, and groupings by outlets, finding that respondents were more likely to agree to cooperate with the survey and maintained their diaries for longer if a product diary was used. The relevance of this evidence to LSMS interview surveys is a matter for conjecture.

Finally, there are implications for the degree of aggregation if the survey is attempting to collect data on physical quantities. Some foods, including many cereals, can be grouped together, and a meaningful total weight can be calculated. However, this is not the case for such items as canned goods, vegetables, or many processed goods; for these items considerable disaggregation is required in order to obtain appropriate units.

In summary, it seems that using drastically shorter questionnaires is likely to be risky and lead to the under-estimation of total consumption. Three hundred items are probably too many and 10 are probably too few but it is difficult to be more precise. The draft module presented in this chapter includes only the approximately 70–100 items that have commonly been used in past LSMS surveys, on the grounds that many more items would increase costs noticeably and that the comparisons to NIPA have not shown huge, systematic biases. Using a much smaller number of items would increase the risk of underestimating total consumption and would certainly decrease analysts’ ability to calculate rough estimates of caloric content (which are sometimes used to calculate the poverty line).

#### *Recall Period*

Each consumption item in the consumption module must be given a recall or reference period. The questionnaire may ask how much rice the household purchased during the previous week, two weeks, or month, or it may ask about the household’s expenditure on clothing during the previous two weeks, month, or year. The recall period is sometimes tied to

a particular event—most commonly the interviewer's last visit. Another option is to have respondents report how much they "usually spend" over a month or a year. While it makes obvious sense to use longer reference periods for items that are rarely purchased and to use shorter periods for high-frequency purchases, this guideline leaves a great deal still to be decided.

**RECALL PERIOD CHOICE.** The choice among recall periods is one of the most important and difficult design issues for the consumption module. It is also an issue that cannot be dealt with in isolation because it interacts with other elements of the module (such as whether expenditures are collected by diary or by interview) and with the survey's design more generally (in particular, whether the design permits multiple visits at least a week apart). The ultimate objective is to obtain a reasonably accurate estimate of the rate of each household's total consumption expenditure over the previous year. There are many different ways to fulfill this aim.

One possibility is for the interviewer to make a single visit to the household during which the respondent is asked to recall how much the household spent during the previous 12 months, either in total or on a list of items. This is likely to lead either to an underestimation of household expenditure (because it is difficult for people to remember their expenditure from so long ago) or to educated guesses (in which respondents estimate their expenditure over the whole year from their current rate of expenditure).

Another alternative is for the interviewer to visit the household many times throughout the year and ask the respondent for details of the household's expenditures over shorter periods. However, if people's memories of their expenditures fade quickly, many visits may be required to ensure that accurate data are collected on high-frequency purchases, and such visits can be prohibitively costly. The diary method was designed to minimize reliance on respondents' memories because the diaries are supposed to be filled out at or near the time when the purchase is made. However, diaries clearly pose special problems when a substantial fraction of the population is illiterate—problems that will be discussed more fully below.

One of the special features of LSMS surveys is the requirement that each survey provide an estimate of annual expenditures at the household level. Most consumption surveys do not make this demand and are

content with estimating aggregates or averages over households, for example, for weights for a consumer price index. If the only problem with reporting were progressive forgetting (the fact that respondents' memories of their expenditures fade as the time since the purchase grows longer) and if there were no other systematic biases (but there are, as will be discussed below), averages could be obtained accurately with short reporting periods. However, this is not adequate for measuring welfare at the level of the individual household. Longer recall periods are better than shorter ones for measuring the distribution of consumption because averaging consumption over many days eliminates the randomness of some of the household's day-to-day purchases that have nothing to do with its standard of living. However, if people find it harder to remember more distant events, longer recall periods will miss more consumption, and lead to downward bias. If short recall periods are used, and if people report accurately, not everyone will purchase every item every day or every week, because many goods can be stored, because many goods that the household consumes regularly do not have to be consumed every day, and because some items of consumption have seasonal patterns. Provided that the survey's fieldwork is spread throughout the year and provided that the respondents' reports are accurate, short reference periods will yield unbiased estimates of the mean for the population. Those households that do not purchase anything during the reference period will be averaged with those who happen to make purchases for several periods, or those interviewed about their consumption during a festival are averaged with those interviewed about non-festival consumption. However, such data (for periods when some households spend nothing while others spend a lot) do not give an adequate picture of the annual consumption of individual households nor of the distribution of consumption across households.

Using a shorter recall period in the consumption module than the period over which living standards are defined in analysis will inevitably cause error in the measurement of living standards. Because adding variance always increases the apparent inequality—effectively a mean-preserving increase in spread—measurement error exaggerates inequality and, if the poverty line is below the mode of the distribution, exaggerates estimates of poverty. In the extreme case, using a single day as the recall period would, while

effectively eliminating bias in the mean, classifies as poor anyone who did not go shopping the previous day. Since the LSMS surveys are at least as concerned with the dispersion of households across the distribution as with means, single-visit consumption modules with very short recall periods should be avoided—except when purchases are known to be evenly spread (say, as a result of rapid inflation that causes people to make frequent, regular purchases). This is not to say that short recall periods and multiple visits can never work well. Several African expenditure surveys use a daily recall period for seven daily visits. In Singapore consumer expenditure surveys use this kind of design in conjunction with diaries (see Silberstein and Scott 1991).

The difficulties that can arise with overly short recall periods are not confined to the estimation of poverty and inequality. If the recall periods for reporting consumption are shorter than the period over which living standards are defined in the analysis, the measurement error in each individual expenditure will be transmitted into the total expenditure estimate, which is the sum of all of the individual expenditures. As a result there will be a nonstandard measurement error bias in the estimation of Engel curves, including calorie Engel curves (Cramer 1969; Bouis and Haddad 1992). If this is not corrected for, estimated Engel elasticities will vary with the length of the recall period (Ghose and Bhattacharya 1993, 1995).

**THEORY AND EVIDENCE IN THE LITERATURE.** People do not forget all events at the same rate, and they can remember some “flashbulb” events (for example, the death of President Kennedy in the United States) in great detail for many years. Yet people generally forget more events the further such events slip into the past. Purchases of consumer goods are no exception to these rules; there is a large body of evidence in industrialized countries on “recall bias” or increasing underestimation as the recall period is increased. (See Neter 1970, Eisenhower and others 1991, Silberstein and Scott 1991, and Sudman and others 1996; for evidence from developing countries see Mahalanobis and Sen 1954 and Ghosh 1953.) Scott and Amenuvegbe (1990), who ran experiments using households from the Ghanaian Living Standards Survey, found that for 13 items frequently purchased by these households, reported expenditures fell at an average of 2.9 percent for every day added to the recall period, within the

first week. Thus expenditures reported seven days later were 87 percent of what they were for a single day. After two weeks reported expenditures were another five percentage points lower. Annual recall based on explicitly normative questions (“How much do you usually spend on xx?”) gave a total 91 percent of that of one-day recall, while annual recall based on ostensibly factual questions (“How much did you spend on xx?”) gave a total 113 percent of the one-day recall figure.

A second important bias caused by memory that affects expenditure estimates is “telescoping,” whereby respondents include in their reports events that happened before the beginning of the recall period. Asked about expenditures during the previous year, respondents may include a car that they bought 13 months ago. (Neter and Waksburg 1964 identified such effects in the U.S. Consumer Expenditure Survey for homeowners’ alterations and repairs.) According to the formal models of Rubin and Baddeley (1989) and Bradburn, Huttenlocher, and Hedges (1994), because people do not remember dates very well, they may remember an event (or expenditure) but be unsure about the date this event occurred. If uncertainty about dates increases as the event recedes, such telescoping errors will cause an increasingly upward bias in the resulting data. The further in the past the event is, the greater is the uncertainty about its date; thus the probability of a less recent event being misplaced into a more recent recall period is higher than the probability of a more recent event being misplaced out of such a recall period.

Short recall periods are more affected by telescoping than by recall bias—leading to overstatement of purchases. As the recall period increases, recall bias becomes more prevalent, resulting in a downward bias (Eisenhower, Mathiowetz, and Morganstein 1991). These effects will work differently for different goods; telescoping has a greater effect for purchases that are highly salient to the respondent—such as durable goods purchases, wedding or funeral expenses, or purchases of grain stock for the year—while high-frequency smaller purchases on such items as food and household supplies are more likely to be forgotten altogether.

Neter and Waksberg (1964) proposed dealing with telescoping using “bounded recall,” a method that appears to be reasonably effective. The first step in this method is to conduct a preliminary interview in

which respondents are asked about the household's expenditures in, say, the previous month. Although these data are not used because they are subject to telescoping bias, the process of noting them provides a record that prevents their being reported again in the first "real" interview—and thus eliminates telescoping. In the second interview, respondents are asked about their purchases since the first interview. Bounded recall is only possible when the interviewer makes at least two well-separated visits to the household. Recall bias is more difficult to deal with, but if respondents are willing and able to keep an accurate diary, both recall bias and telescoping are eliminated.

Cognitive psychology and questionnaire "think aloud" interviews are yielding insights into how people answer questions. It seems that as long as an interview is well conducted, most respondents do their best to provide accurate and truthful answers. At the same time, most respondents also try to minimize the effort they must make to respond to the interviewer's questions, thus acting as "cognitive misers." As a result, they may switch their tactics for answering expenditure questions as the difficulty of the task increases. Over short recall periods or for rare but important events in their lives, their answers are based on counting; they recollect individual events and then add them up. Over long periods, for unimportant events that happen frequently, or for aggregates containing large numbers of items, they often resort to a more approximate way of estimating their answer: estimating the frequency of the occurrence and multiplying this number by the length of the reference period. The frequency that they choose may or may not be accurate; it may, for example, be overweighted toward their current or recent behavior, with the respondent ignoring or giving inadequate weight to exceptional events. Sudman, Bradburn, and Schwarz (1996) stated that their "tentative finding is that estimation is unbiased but counting methods, although reducing variance, may be biased either up, for short time periods, or down, for long periods." In this sense, respondents' estimation strategies can be thought of as an alternative to diaries as a means of dealing with recall and telescoping biases. Indeed, there is some evidence from consumer expenditure surveys in Canada (McWhinney and Champion 1974) that diaries and interviews with annual recall periods gave closely similar results.

The U.S. Consumer Expenditure Survey, a survey that could be seen by statistical offices in developing

countries as a model to emulate, has changed its design to accommodate the findings in the literature (see Jacobs and Shipp 1993). Although the Consumer Expenditure Survey is fielded in an industrialized country and although its primary function is to collect weights for the price index rather than to monitor welfare, it collects detailed and aggregated consumption measures, and its experience is relevant for designers of LSMS surveys. There is good documentation on this survey's experience with telescoping, recall bias, seasonality, and the advantages and disadvantages of diaries compared to interviews.

In Consumer Expenditure Surveys up to 1960–61, expenditures were obtained by interview using questions with an annual recall period. (An interesting feature of this methodology was a "balancing" procedure whereby a household was revisited for cross-checks if the household's reported expenditures and reported income differed by more than a prespecified limit.) However, in the decade before the next survey (in 1972–73) came the work on telescoping by Neter and Waksburg (1964), as well as influential experiments on diaries and interviews by Sudman and Ferber (1971). The annual recall period was abandoned, as was the "balancing" procedure—which was considered arbitrary and unworkable in the absence of an annual recall period. Available literature does not provide any evidence that the annual recall procedures were unsatisfactory or that the balancing procedure failed to work—only that such techniques had been superseded.

In the current design of the Consumer Expenditure Survey, introduced in 1980, one set of households keeps diaries that cover expenditures on food and minor household items (mostly grocery items), and a different set of households is interviewed on five separate occasions. The data from the first interview, which uses the recall period of the previous month, are used only to eliminate telescoping bias. At each of four subsequent quarterly interviews, households are asked to recall their expenditures during the previous three months. The evidence from these surveys is entirely consistent with previous evidence about the significance of telescoping bias. According to Silberstein (1990) the (discarded) rates of expenditure from the first interviews are much higher than those from subsequent interviews; for clothing the first interview totals are 40 percent higher than the average of the subsequent four interviews. However, there is

also “internal” telescoping or recall bias in the data from the subsequent interviews, with respondents consistently reporting higher expenditures in the most recent (third) month of each quarter than in the two months preceding it. There are also pronounced seasonal effects in the reported expenditures in the Consumer Expenditure Survey—especially increased consumption associated with the year-end holidays (Silberstein and Scott 1992). The U.S. Bureau of Labor Statistics doubles the size of the sample over the holiday period to deal with these and other (largely non-response) effects; the effectiveness of this measure is not clear.

**EXPERIENCE FROM PAST LSMS SURVEYS.** Past LSMS surveys have used a range of recall periods for consumption items, depending on both the item and the survey. For food purchases the ongoing Jamaican survey uses 7-day and 30-day recall periods. In South Africa in 1993 respondents were asked whether they bought each food item on a weekly or monthly basis, and were then asked to report their purchases during the last such period. The survey in Ecuador (1994) took a similar approach in that the respondent chose the recall period for food items. The recall period was one week in the Kyrgyz Republic (1993), Nicaragua (1993), and Russia (1993–94) surveys and two weeks in Brazil (1996). In China (1994; Hebei and Liaoning provinces only) the period was specified simply as “1994.” In many surveys nonfood items have often been separated into two categories: high-frequency or “daily” items and “occasional” items. Daily items have a short recall period—perhaps a week or two—whereas occasional items may have recall periods of one month, three months, six months, or a year. For nonfood items, some surveys have two recall periods; the Jamaica survey uses a month and a year. Other surveys sort different items into different single recall periods; expenditures on soap may be reported on a monthly basis, clothing on a quarterly basis, and vacations on an annual basis.

One design is frequently thought of as an LSMS standard, in part because it was used in several of the earliest and most widely analyzed LSMS surveys. In this protocol, respondents are asked whether the household has consumed a particular food item during the past year. Each respondent who answers “yes” is asked a series of follow-up questions. First, the respondent is asked whether the household purchased

any of the foodstuffs on the list. If the answer is “yes,” a further question is asked about the value of any purchases that the household made since the interviewer’s last visit (a time that in the prototypical LSMS fieldwork plan is 14 days). Second, the respondent is asked in how many months of the year the household purchased the food item, how often it purchased the item in each of those months, and how much it usually spent each time. Data on the value of home-produced food are collected in a separate set of questions that ask how often the home-produced food is consumed; the recall period for these questions has varied from country to country in previous surveys, ranging from “each time the home-produced food is consumed” to each day to a typical month.

This design makes it possible to compute two different estimates of the monthly rate of expenditure for each food item. The “last-visit” measure is zero if no purchases were reported; otherwise it is the amount reported since the last visit divided by the number of days since the last visit and multiplied by the number of days in an average month (365 over 12). The “usual-month” measure is zero if nothing was purchased in the previous year; otherwise it is the reported usual monthly expenditure multiplied by the number of months in which purchases were made and divided by 12. Most analysts of the data, whether constructing poverty profiles or conducting research, have used the usual-month figures.

In some surveys the respondent has been offered only a single recall period for some nonfood expenditures—a week, the time since the interviewer’s last visit, a month, or a year, depending on the presumed frequency of expenditure. Frequently purchased nonfood items such as newspapers and tobacco are usually collected with a recall period of the previous seven days or the time since the interviewer’s last visit. However, for a substantial number of items, a dual procedure similar to the procedure used for food has often been followed, with a last-visit measure constructed as before (although the alternative is now expenditures “in the last year” rather than usual monthly expenditure). In this case a monthly estimate can be constructed by dividing the response by 12.

The last-visit measure can be thought of as an explicitly bounded measure that elicits from the respondent an answer based on his or her recall and counting of events, while the “usual-month” question is an attempt to elicit from the respondent an answer

based on a rate or frequency. Of course, these are generous interpretations. In the standard protocol, no consumption questions are asked at the first visit. Thus, although the first visit may be fixed in the respondent's memory, it is not clear that this will help to reduce telescoping errors at the second visit. The respondent may remember the previous visit of the interviewer very well but still be unable to recall whether his or her trip to the market occurred before or after that visit.

There is also a serious question about whether the usual-month recall is likely to be independent of the last-visit recall. If respondents are "cognitive misers" casting about for the easiest way to estimate the "usual" rate, the answer to the "last-visit" question will give them one. Omitting the "last-visit" question might not even solve the problem, since the respondent's answer to the frequency question could also be unduly influenced by his or her recent behavior.

In background work for this chapter, data from the LSMS surveys in Côte d'Ivoire (1986), Ghana (1988), Pakistan (1991), and Vietnam (1992–93) were used to compare the consumption estimates from the different recall periods used in the consumption modules. The aim was to look for evidence that reported mean expenditure rates decline with the length of the recall period and to check whether the same is true for the dispersion of the estimates.

For purchases of food (excluding the value of home production), the estimates are very similar for the last-visit and usual-month methods of calculating monthly food expenditures. Where there are differences, they do not conform to the expected pattern in which both mean and dispersion are lower in the usual-month figure. Indeed, the differences are most marked in Côte d'Ivoire, where the last-visit measures have a lower mean (by 5 percent) and median (by 8 percent) than the usual-month estimates, and the measures of dispersion are very close. For Ghana the two sets of numbers are effectively identical, a finding that extends to the complete distributions so that no poverty or inequality measure would be different if one kind of data were used to calculate it rather than the other. In Vietnam the last-visit measures are slightly lower than the usual-month measures, and the dispersion in the latter is perceptibly lower. A close inspection of the details revealed that there are fewer very low reports in the usual-month data than in the last-visit data. (The Pakistan data cannot be used for this comparison due to the problems discussed above.)

For nonfood items, for which the frequency of purchases is generally lower than for food, the differences between the two measures are more marked. For the means, the usual-month measures are lower than the last-visit measures for Côte d'Ivoire, Pakistan, and Ghana but not for Vietnam. The dispersions are lower for the usual-month measures than for the last-visit measures in all four surveys. Except for Vietnam, the decline in means is consistent with the syndrome of increasing forgetfulness over a longer period, and all the dispersions are consistent with the view that the last-visit (two week) measure of these items is too short to give an accurate measure of households' annual standards of living.

When food and nonfood are put together with other items to give a consumption aggregate, dispersion of total per capita household expenditure from the usual-month data is less than for the last-visit data, but the difference is not very marked, given the fact that the item with the biggest drop in dispersion—nonfoods—is a relatively small share of most households' budgets. Nevertheless, the differences are large enough that the headcount ratio measures of poverty will all be lower when usual-month data are used rather than last-visit data.

These results can be seen as encouraging, since they show only slight sensitivity to the choice between two of the most obvious reporting periods. An optimistic interpretation is that food expenditures are frequent and stable enough that the last-visit and usual-month estimates are similar. While this is not true for infrequent expenditures for which the choice of reporting period affects both bias and variance, the total expenditure on such items is usually too small for their net effect on total consumption to be very large.

There are some caveats to this finding. If respondents forget about their food purchases at the sort of rates suggested by Scott and Amenuvegbe (1990), the last-visit measures may be substantially underestimated. In addition, the consistency between the usual-month and last-visit measures for food may simply reflect households using their responses to the last-visit questions to guide their answers to the usual-month questions, which would not then constitute any sort of independent check on the validity of either measure. And there are some nonfood items in these surveys (mostly frequently purchased items) for which there was only a single recall period. While it might be hoped that these nonfood items would not be sensi-

tive to the choice of recall period, the inclusion of these items in the consumption totals mutes the effects of using different recall periods.

**OPTIONS FOR FUTURE SURVEYS.** There are no definitive answers about the optimal recall period. In the meantime, however, surveys must be designed, so we provide a brief discussion of the various options together with some suggestions.

If a survey is meant to be comparable to another survey, it makes sense to use the same recall periods that were used in the other survey, provided that the previous survey conformed with best practice standards as outlined in this section of the chapter.

Beyond this consideration, there are two main routes to follow. The first is essentially the “status quo” of the design of most recent LSMS surveys, while the second is a more extensive revision based on the supposition that the current design is unsatisfactory. The preference of the authors of this chapter (which is reflected in the draft questionnaire) is to make only minor modifications to the status quo and to experiment with the components of any revision before putting it into practice.

The “status quo” design is to use two recall periods—one that is the amount of time since the interviewer’s last visit and the other that is for “usual” expenditures. As already argued, the first yields estimates that have a minimal amount of telescoping, while the second is a calculated, unbiased estimate. The evidence of the comparisons with NIPA data can also be cited; these do not show the gross and systematic under-estimation of expenditures that might be expected if the recall data missed a very large fraction of expenditures. Finally, an important characteristic of LSMS surveys has been their coverage of many topics, which would be threatened by including a vastly more extensive consumption module. Consumption is already one of the longest and most expensive modules of any LSMS questionnaire, and to extend it further (even if the resources were available to do so) would inevitably crowd out other important topics. Nevertheless, these are not strong arguments. The NIPA comparisons are weak, the consistency checks in the surveys are capable of many interpretations (not all of which support the validity of the data), the last-visit recall period contains no bounding questions, and arguments in terms of expediency are much weaker if the data are of uniformly poor qual-

ity. Moreover, the two-visit structure is being used increasingly less frequently; when this structure is not used the last-visit question has to be replaced by a question with an unbounded recall—increasing the chances of telescoping.

The other point of view is that the LSMS standard questions are far from best practice. One of our reviewers, the late Chris Scott, said, “the use of the question, ‘How much do you usually spend on mangoes?’ appears to fall far outside of reasonable best practice.” Scott, who had extensive experience carrying out consumption surveys in Africa and elsewhere, believed that only a more in-depth approach would yield adequate data. According to Scott, respondents should be interviewed several times, separated by the most accurate recall period, perhaps as little as a day. The number of interviews should be as many as are needed to cover the reference period (say, a week or two), with a bounding interview at the outset. Some or all of these interviews can be replaced by diary-keeping by the respondents themselves, by proxy recordkeepers in the household, or by interviewers who complete the diaries with help from the respondents, an alternative that blurs the line between diary and interview procedures. There is some evidence (reviewed below) that diaries can capture some expenditures that may be missed in interviews, and with a sufficient number of such diaries, the substantial additional cost of conducting interviews can be translated into high-quality data. Nevertheless, even with a two-week reference period (requiring 10 or more interviews), seasonality is not captured, nor are other fluctuations in consumption over the year. It is conceivable that usual-month responses capture some of this variation, but this is probably little more than a hope.

Every survey should have a budget for experimentation. Questions should occasionally (if not regularly) be subjected to cognitive laboratory techniques and revised and updated in the light of the results. Even more importantly, the extensive interviewing and diary techniques, as advocated by Scott, urgently need to be compared against the standard last-visit and usual-month responses, preferably using randomly selected subsets of households within the same survey.

Finally, it cannot be overemphasized that the wordings of the recall periods must be unambiguous and well understood by interviewers. Wordings such as

“since my last visit, two weeks ago” are obviously ambiguous if the visit did not occur exactly 14 days ago. There are also possibilities for confusion in recording the units of a purchase. Where quantity or price data are sought, it must be clear what units the questions refer to. Obtaining the units for quantities purchased may work better than obtaining the units for prices or unit values; this is reflected in the draft consumption module.

#### *Multiple Visits Throughout the Year*

In most past LSMS surveys, primary sampling units and the households within these units have been visited on two occasions two weeks apart, with the consumption data collected only during the second visit. Thus there is a single record of consumption over whatever recall period is selected. This has also necessarily been the case in the increasing number of LSMS surveys where only a single visit is made to the primary sampling unit. Even if the surveys were to adopt a more intensive program of multiple interviews, consumption data would still be collected over only a relatively brief period—say, a week or a month. No proposed design of the LSMS consumption module would capture variations in household consumption over a whole year; therefore, the consumption data currently collected may not reflect the annual consumption flows in which analysts are fundamentally interested. To collect better data it would be necessary to revisit households on several occasions throughout the year, collecting consumption data during each visit. Such data could be used to increase the accuracy of the consumption aggregates, allowing for variation in households’ consumption over time; the data could also be used for a number of analytical exercises.

One reason for multiple visits throughout the year would be seasonality. The collection of agricultural data and agricultural income typically requires that survey interviewers visit households (farms) in different seasons of the year. Farm incomes are seasonal, and it may not be possible for a respondent to remember all of the transactions that went into a calculation of net income many months afterwards. In consequence, it is widely believed—and suggested in Chapter 19 of this book—that accurate farm surveys require multiple seasonal visits. The concern in the current chapter is with consumption, but it is conceivable that consumption expenditures—like incomes—vary systematically with the seasons, so that if respondents’ mem-

ories are a problem, accurate estimation of annual consumption flows will require multiple, seasonal visits just as does the accurate estimation of income.

The seasonality of consumption does not necessarily follow from the seasonality of income, and may have a quite different seasonal pattern. In theory, farm households have strong incentives to untie their consumption from the seasonal patterns of their incomes. Harvests may generate income for only a few weeks or months, while consumption has to be maintained throughout the year. It is not necessary to accept the permanent income or life-cycle hypothesis of consumption to believe that households can smooth their consumption over the year. Even without access to credit markets, farmers can store some output or save some income from the harvest to support consumption throughout the rest of the year. Of course, there will still be seasonal patterns in farmers’ consumption; festivals such as Tet and Christmas are associated with higher-than-normal expenditures. And because storage is costly, prices will generally be higher just before the harvest than just after it, with some effect on consumption. Nevertheless, seasonal variation in consumption is neither closely tied to, nor less than, seasonal variations in income. Reliable evidence on the relative seasonal smoothness of consumption comes from Paxson (1993). Rice farmers in Thailand who double-crop (with irrigation) have quite different seasonal patterns of income than do farmers with only one crop, yet their consumption patterns are almost identical and exhibit little seasonal variation. The same is true when one compares farm and nonfarm households or farmers in different agroclimatic zones.

For the purposes of this chapter, an attempt has been made to look for seasonal patterns in consumption data from three LSMS surveys. Cross-sectional surveys are not well suited to this task, since what is required is not a large number of households observed throughout a single year but rather a large number of years over which seasonal patterns can be established. Nevertheless, the consumption totals for Ghana, Côte d’Ivoire, and Vietnam were examined for differences across months. In any one survey for a single year, there tended to be significant differences in the consumption total from one month to another—differences probably not driven by seasonal patterns in consumption. (In the Côte d’Ivoire and Ghana surveys, which include data for more than one year, the monthly patterns in the consumption total are quite

different across the survey years.) Some of these differences—such as those caused by Tet in Vietnam—are easy to explain, but others are not, even though their effects are sometimes considerable. The progression of the survey teams through the country may generate some variation as they move from poorer to richer villages. Random measurement error is also likely to be an important factor. Nevertheless, this limited analysis provides no grounds for supposing that consumption would be better measured by including multiple, seasonal visits to households in LSMS surveys in the future.

However, there are other reasons why multiple visits might be useful. The average of two consumption totals, each for a two-week recall, will give a better, lower-variance estimate of longer-term consumption. More radically, Scott (1992) and Central Statistical Office of Zambia (1995) advocate using the correlation between consumption across multiple visits to correct measures of inequality to bring them closer to what would have been measured had it been possible to collect consumption data over a full year for each household. The idea is as follows. Suppose that analysts are interested in annual consumption measured at a monthly rate but the only observations they have are of consumption over the previous month. At one extreme, each household may consume the same amount in each month so that the monthly totals are correct and can be correctly used to give a measure of dispersion over households. At the other extreme, suppose that each household's consumption is uncorrelated from one month to the next. The "last month" totals are then correct on average, but their dispersion over households is larger than the dispersion in which analysts are interested because it has a "within-household" component in addition to the "between-household" component with which analysts are concerned. With multiple observations for at least some households and under some reasonable assumptions, it is possible to estimate the size of the within-household dispersion and to correct the measures of the total. The Central Statistical Office of Zambia (1995) uses such a procedure to correct dispersion measures for a Zambian survey; the technique could usefully be applied elsewhere.

As always, measurement error will add dispersion to measured consumption, so that if measurement error is random, the dispersion it causes will be in addition to the genuine dispersion that comes from

the intrahousehold and interhousehold components of dispersion. Although this is not explicitly allowed for in Scott's formulation, his corrections should still lead to a better estimate of dispersion. With repeated observations of the same households, there are a number of techniques that would make it possible to assess the size of measurement errors (see particularly Griliches and Hausman 1986).

Finally, in some circumstances multiple consumption measures can be useful at the analytical stage. Multiple visits generate a type of longitudinal or panel data that can be useful for studying changes over time and for sorting out cross-sectional variation from time-series variation, as in Scott's work above. However, if the visits are separated by only a few months, the changes that have occurred may not be large enough to be interesting, and the measured changes are likely to be dominated by measurement error.

The recommendations made here are again tempered by the increased costs of extending the consumption module of a typical LSMS survey, in terms of both money and the consequences for the rest of the survey. Because consumption is smoothed within the year, measuring it over two weeks or a month may yield a sufficiently accurate picture of annual consumption to make it not worth incurring the cost of adding yet more visits. However, if the detailed agricultural module is included in the survey with multiple visits during different seasons, there would be little additional cost involved in collecting at least some consumption data at each visit. Such cases aside, including multiple visits throughout the year is probably not the highest priority for improving the typical LSMS survey.

#### *Imputing Values*

In nearly all LSMS surveys, calculating a comprehensive measure of consumption will require at least some imputations. Not all consumption is obtained through market purchases; if analysts want to calculate consumption in monetary units, they must find some way of pricing its unmarketed components. In many of the poorest countries, and especially for the poorest people, a large share of food comes from home production or from hunting, fishing, or collecting wild foodstuffs. These imputations for food are likely to be those that are most important for the totals. However, there are a number of other commodities obtained by the house-

hold's nonmarket labor, such as homemade clothes or wood and water fetched by children or women. Household members often receive gifts or in-kind payments that need to be priced before they are added in to the consumption total. And it is necessary to obtain information on the consumption that comes from using durable goods. For nondurable goods—even those that are partly durable—it is probably safe to assume that consumption and purchases are one and the same. However, for large durable goods that are expensive and last for many years, such as houses, cars, or bicycles, it is important to try to make some adjustments. This subsection reviews the data required to make these imputations.

It should first be noted that imputation is an inherently difficult and error-ridden process. Imputation is likely to work best where there is relatively little need for it—when the economy is highly monetized but there is a relatively small amount of own-production (such as vegetable gardens) involving goods that have clear market equivalents. Imputation works badly in economies in which a large share of transactions do not pass through the market. LSMS procedures for estimating welfare stem from a theory of a consumer with well-defined preferences operating in a market where prices are well defined and unaffected by the agent's behavior. Where these markets do not exist, analysts are in effect imposing an accounting framework on the physical data, a framework of dubious relevance to the lives of the people being studied.

Food that is either home-produced or received as gifts or payment in kind has been the most important imputed item in LSMS surveys to date. In principle, the calculations are straightforward. The respondent is asked to report the values of any home-produced food items consumed by the household during the reference period, and the sum of these values is added to the consumption total. Given the seasonality of production, the recall period probably has to be a year, or at least a typical month over the last year. It may be possible to do better than this when there is a multiple-visit agricultural module in the survey. However, the major difficulties are with valuation, since the respondent is being asked a purely hypothetical question about the sale or purchase of an item that is rarely traded or that may have been traded some time ago.

The draft module (presented in Volume 3) recommends collecting data on physical quantities of goods

consumed, since these are magnitudes that the respondent observes, at least in principle. The value or price of these quantities can be obtained in several ways. Farmgate prices, defined as what the household could get for its production, set a lower bound on valuation, since it is usually presumed that consumption is evidence that the good is valued beyond what it would fetch. Market prices, by contrast, are likely to be too high because they include transport and distribution margins and because the commodity traded is often of higher quality than its home-grown counterpart. However, once the quantity has been obtained, the respondent could be asked to report one or both of these two prices or simply to estimate the value of the commodity directly. Some degree of cross-checking is possible from the quantities and prices of purchases reported in the agricultural module or from the prices gathered in the community questionnaire.

In some circumstances it may be possible (or important) to carry out a similar exercise for nonfood items. Clothes and furniture are often made at home, and household labor is used to collect firewood, dung, or water. These items have usually been omitted from past LSMS surveys, probably because of difficulties in valuation. One danger is that the welfare of poor households might be overstated by using inappropriate prices or wages to value their production or their labor. For someone with no other employment opportunities who ekes out a living gathering firewood or coal, it would be adding insult to injury to impute a high standard of living to them by valuing their time in terms of the market wage in a formal sector to which they have no access. In some circumstances the wage data from the community questionnaire may be a better basis for imputations based on the value of time, but these data do not eliminate the dangers inherent in the procedure.

It is incorrect to compute an extended concept of consumption by adding “expenditure” on leisure (in other words, the value of leisure at the market wage) to total consumption. This “full-income” concept has its uses, but it is a nominal measure and, like other nominal measures, must be converted to real terms before being compared across households or individuals. The problem here is that even if everyone in the comparison faces the same (or similar) prices for goods, they do not face the same price of leisure, because wage rates differ. As a result, before full income can be used as an indicator of welfare, it must

be deflated by some price index that includes the price of leisure. Alternatively, if the value of leisure is to be added into the value of total expenditure, the same wage rate should be used to value everyone's leisure.

For durable goods, the consumption flow is best thought of as a rental equivalent or "user cost." This has two components: the opportunity cost of the funds tied up in the good (that could be realized through its resale) and the value of the physical depreciation of the good (through use of the good or passage of time). To estimate these magnitudes, some measures of depreciation and current value are needed. Perhaps the simplest way to obtain this information, at least for goods purchased in the previous five to ten years, is to ask respondents when they purchased the good and how much it cost at that time. These are both factual matters and will often be clearly remembered, at least in the case of large, important items. Provided that the good has been available for some time and purchases have been made relatively evenly over time, an estimate of the average lifetime of the good can be obtained by doubling the average age over all similar goods for all households in the survey. Once this approximate lifetime is known, the depreciated value of the good can be estimated given its age and original value, which is then used to calculate the first component of the user cost. It is also possible to ask respondents for direct reports of the current market value of the used durable, though there is no evidence on the accuracy of such (hypothetical) reports, and field tests are likely to be useful.

For housing, the largest of the durable goods, the imputation approach again starts from the rental equivalent. Unlike the value of most other durable goods, rents can sometimes be observed directly, and these are the correct numbers to add into the consumption aggregate. For households that do not report rents, the standard procedure is to impute a rent based on the characteristics of the house, as reported in the housing module. This is typically done through "hedonic" regressions in which reported rent is regressed on the house's characteristics (such as size, number of rooms, construction material, and location) and the results are used to calculate rents for other properties where rents are not reported. The credibility of these regressions is compromised if only a small fraction of the sample reports rents and, more generally, if those who report rents are unrepresentative of the population as a whole. While it is possible to make mechan-

ical corrections for the selection, these corrections usually require arbitrary and untestable assumptions that further compromise the credibility of the process. This is a difficult area. In general, survey analysts should make sure that indefensible imputations are not dominating welfare comparisons. The data required for rent imputations are gathered in the housing module (and to some extent in the community questionnaire) and therefore are not discussed further in this chapter.

A number of imputations come from other modules in the survey. The employment module gathers information on in-kind income provided by employers, including (free or subsidized) transport to and from work, food at work, and housing.

#### *Respondents for the Consumption Module*

Most LSMS surveys have interviewed a single respondent for the whole of the consumption module or for each part of it. The household is asked to determine the "best informed individual" who will respond to questions. This has the appealing feature of not prejudging the division of labor in a household, either by gender or by age, as would be the case if it were assumed that the wife does the shopping (or if any similar assumption were made). In many—perhaps most—countries, the single-respondent approach works well. In particular, this approach is satisfactory where food is a large share of the budget, where there is a common cooking pot, and where most of the household resources are pooled.

Even within resource-pooling households, it may be useful to have multiple respondents for some expenditures or different household members reporting on different categories of expenditure. While the person who does most of the shopping for food will know about this large share of the household's budget, another large share could be most accurately reported by the person who pays the housing and utility bills (who may or may not be the shopper). And there are other expenditures of which no one single person may have a very accurate picture. Individuals may not know how much "walking around" money other household members have, much less how they spend it—whether on bus fares, meals away from home, newspapers, tobacco, alcohol, or entertainment. Also, there may be larger items, such as clothing, that individuals purchase without any other household member knowing how much was spent. This is partic-

ularly true where several adults live in a household, each contributing some amount of their income to the household's joint expenses and reserving the remainder for their own use. For example, in a household with a mother and grown sons, the mother may pay for all the household expenses, including food and utilities, using her pension and the regular monthly sums given to her by her sons. However, she may have little or no idea about her sons' incomes or other expenditures, which could account for most of the household's total income and outlay. In this situation, no one person can give an accurate report of the household's income and expenditure, nor is the "household" really the relevant unit for analysis.

The LSMS has little experience in procedures for dealing with these situations. The literature discussed in the next subsection describes cases where each adult member of the household kept a diary of at least some categories of expenditure. It is possible in principle to interview each member of a household about at least some expenditures, such as those paid for by "walking-around" money. A more ambitious prospect would be to try to record incomings and outgoings for each member of the household who spends money. This would probably be prohibitively expensive for a general multitopic survey, although as always there are potential benefits from conducting experimental work either on a few households or within a special survey. Multiperson accounts are likely to provide a fascinating picture of how intrahousehold transfers of resources take place, who gets what, and what makes a group of people function as a household. This is an important research area but not a prime candidate for immediate incorporation into standard LSMS surveys.

#### *Diaries Versus Oral Interviews*

The use of consumption diaries, in which households are asked to record their purchases as soon as they make them, is common in full-fledged single-purpose consumption surveys. The ideal diary would yield a record of each purchase immediately after it takes place, thus eliminating the need for respondents to rely on their memories and removing any associated errors, including telescoping. The diaries may be kept by a single respondent or by several or all members of the household, so they can potentially help resolve the question of who is the best respondent for the whole household while simultaneously yielding information on intrahousehold allocations. Diaries may be organ-

ized on a product basis, on an outlet basis, or on a purchase basis, and the forms can be designed to allow a large degree of prompting (for example, by listing many types of products) without the associated tedium of a long interview.

The use of diaries has some practical implications; the most obvious is that a person filling out a diary must be literate. Nevertheless, the diary approach has been used in household expenditure surveys in a large number of countries where literacy is not universal. This has been achieved by having the most literate member of the household (sometimes a child) help the one who does the purchasing fill out the diary, or by having the interviewer visit the household frequently—perhaps even daily—to help the household fill it out (Blaizeau 1998). In such cases, the distinction between a diary and an oral interview becomes blurred. This blurring occurs even in literate households, in which the members of the household may either forget to fill out the diary or get tired of doing so; in the U.S. Consumer Expenditure Survey substantial numbers of diaries are completed by the interviewer at the time of collection based on the respondent's memory. When collecting a diary, the interviewer examines it briefly and, if it appears to be incomplete, tries to prompt the respondent to fill it out more completely—essentially transforming the situation into an interview. To the extent that these "diaries" rely on respondents' memories, telescoping and recall bias again become potential problems.

A second logistical issue is that the diary must be both left with the household and picked up after its completion. If the diary period is relatively short—say, a week or two—this does not necessarily pose a problem in an LSMS survey, since the completion of the whole questionnaire typically involves multiple visits by the interviewer to the household so that each member can be interviewed and the length of each interview can be kept reasonable. However, leaving a diary with a household for a long period of time, such as a month or a quarter, would be more difficult and would not be possible within the current design of most LSMS surveys. Thus diaries are only usable for studying items for which a relatively short recall is appropriate—and not for a large proportion of non-food expenditure.

Third, the use of a diary alters interviewing. A diary reduces the amount of time that the interviewer has to spend interviewing households that fill out the

diary completely. However, using a diary may increase the time that the interviewer must spend traveling, since it requires an extra trip to the household to collect it; considerable time may also be spent helping illiterate households fill out diaries.

Using a diary also shifts the burden of response on to respondents. The effect of this shift is unclear; some survey statisticians speculate that households may enjoy the novelty of filling out diaries, and some think that diaries allow households to participate in generating survey data at a time and place that are convenient for them.

If these two speculations are true, then diaries may reduce the burden of surveys on households, making the households more willing to participate in surveys. Yet evidence from industrialized countries shows that keeping diaries is likely to deter households from participating in surveys and that the burden of keeping a diary causes respondents to drop out of a survey over time. There is also evidence that the rate of reporting declines with time, so that, in two-week diaries, more consumption is recorded in the first week than in the second. This has also been true for the Consumer Expenditure Survey in the United States, and for surveys in seven West African countries, as documented by Blaizeau (1998). In the 1995 and 1996 income and expenditure surveys in Belarus, the expenditures recorded in the second week were about 15 percent lower than those recorded in the first week of diary keeping (Martini and Ivanova 1996). In Armenia the diary was kept for four weeks, and the downward trend continued over this longer span. The second week's expenditures on food were 26 percent lower than those of the first, the third week's were 35 percent lower than those of the first, and the fourth week's were 40 percent lower than those of the first (calculations done for this chapter). This could be caused by respondent fatigue or by the fact that the novelty of diary keeping wears off. Also, the fact that they are keeping a diary may cause people to spend more or to shift their expenditures forward into the diary keeping period. Keeping a diary may cause people to think more about their consumption and perhaps take the opportunity to buy some items that they needed anyway. To the extent that diaries are not filled out every day, there is also scope for telescoping and recall errors within the diary period.

Having several members of the same household each keep a diary is also an attractive option in some

circumstances, and there is considerable variation in practice among different kinds of surveys. For example, in the U.S. Consumer Expenditure Survey there is a single household diary, while in the British Family Expenditure Survey all adults in the sample keep individual diaries. There is also evidence on the use of multiple diaries from Hong Kong (Grootaert 1986), Papua New Guinea (Gibson 1998), and surveys in Benin, Burkina Faso, Côte d'Ivoire, Mali, Niger, Senegal, and Togo (Blaizeau 1997). This literature finds that multiple diaries can be useful for obtaining records of expenditures that would otherwise be missed, but that because it is difficult to get all household members to cooperate, attempting to collect multiple diaries can reduce response rates. It is also clear that for some household members—certainly children and perhaps some of the elderly—proxy or household-level reports are more accurate than individual diaries.

To the existing literature can now be added evidence from experiments in three former Soviet republics: Latvia, Armenia, and Ukraine. In the Latvian experiment (analyzed by Scott and Okrasa 1998), a nationally representative set of 300 households was given oral interviews and asked to keep diaries covering a comprehensive lists of foodstuffs as well as a selection of nonfood items. In half of the sample, the diaries were administered first; in the other half, the interviews were done first. The order of the instruments had virtually no effect on the results. Overall, food expenditures were about 46 percent higher for the diary than for the interview; the coefficients of variation were quite similar. This pattern also held for 13 of the 15 subgroups of food and generally for the quantities as well as for the expenditures. For nonfood, the results were much more mixed. For one of the four categories, the expenditures reported in the diaries were significantly greater than those reported in the interviews, for two of the categories, the opposite was the case, and in the fourth category, the differences were not significant.

In Armenia, as part of nationwide survey, one-quarter of the households in each cluster were given diaries for 30 days and interviews with 30-day recall periods were conducted with the remaining three-quarters of the households. The only diary data available to the authors in usable form are from the food sections of the diary, so the comparisons done for this chapter were limited to those items. Total expenditures

for food recorded in the diary were about one-third higher than those resulting from the interview, a result similar to that in Latvia. However, this pattern did not hold as strongly for the subgroups of food. Of the 15 groups (which each included between 1 and 19 items), the diary yielded significantly higher expenditures than the interviews in eight cases, significantly lower expenditures in two, and not significantly different expenditures in the other four cases. There is no noticeable pattern to these results in relation to the mean expenditure, the number of items in the subgroup, or the average frequency of purchases. The coefficients of variation for the diary were twice as large as for the oral interviews. The same pattern of results generally held when the whole sample was divided into rural and urban areas, although the differences were slightly larger in rural areas. The diary estimates for total food expenditure were 43 percent higher than the interview estimates in rural areas and 31 percent higher in urban areas.

In Ukraine a diary was administered to about 500 households in selected locations as an experiment in conjunction with a national survey that used interviews. The diary used a recall period of four weeks for all items. The oral interview used a recall period of two weeks for food items and four weeks for nonfood items. The lists of items in the two instruments were not the same. The comparisons made here were limited to those categories that were either identical in diaries and interviews or that could be clearly mapped into each other. For instance, "butter" was an item listed as such on both lists, but the interview list contained a single question about "smoked sausage and other smoked meats" while the diary had separate items for "smoked sausage" and "other smoked meats." Thus, in this experiment, there are two possible effects that will work in opposite directions. On the one hand, the diaries might yield higher expenditures than the interviews, partly because of the greater disaggregation of the items on the diary list. However, the shorter recall period in the interview may have caused the interviews to yield higher expenditure numbers due to the greater proportional influence of any telescoping. Moreover, if diaries are kept less rigorously as time goes on, the shorter recall period of the interview may mean that higher numbers are reported in the interviews than in the diaries. These effects cannot be disentangled because data available to the authors contain only the subtotals for each item

recorded in the diary, rather than each individual purchase of it.

In the Ukrainian experiment the expenditures for the subtotal of food items that could be matched between the two instruments are 10 percent lower in the diaries than in the interviews—a difference that is significant only at the 10 percent level of confidence. In only 3 of the 11 categories are expenditures significantly different between the diary and the interview. The category that covers bread and flour accounts for about two-thirds of the total differences found in the food subtotals. For the subtotal of nonfood items that could be matched between the two questionnaires, the diary subtotals were 7 percent lower than those for the interviews, a difference that was not significant.

There are several features of these three experiments that limit the extent to which they can be generalized to other contexts and countries. The populations of Latvia, Armenia, and Ukraine are literate, so these experiments cannot indicate what would happen if diaries were used in largely illiterate survey populations. All three countries have long traditions of household expenditure surveys that use diaries. In Latvia and Armenia the experiments were carried out by the statistical office using regular interviews, so the experiments used experienced staff who were thoroughly familiar with the procedure. This was not the case in Ukraine, where the experiments were done by a private survey organization; this difference may account for some of the disparity between the two sets of results.

Past LSMS surveys have made less use of diaries than they could have, and perhaps more use should be made of them in the future. Certainly an ambitious statistical office in a developing country, looking to industrialized countries for inspiration, would use diaries. It is less clear whether, even in wealthy countries, there is indisputable evidence of the superiority of diaries to interviews that might justify such a decision. While diaries would produce better results if they were used in ideal circumstances, it is unclear whether their practical superiority over interviews has ever been convincingly demonstrated. And given the significant illiteracy rates in many poorer countries, the argument for switching to diaries is weakened even further. While there is no doubt that diaries can be used in situations where interviewers make many visits to the household to help them remember their purchases and complete the diaries, this is closer to diary-

keeping by the interviewer than diary-keeping by the respondent. In countries such as Armenia and Latvia, where diaries have been routinely used in the past, there is every reason to incorporate diaries into LSMS-type surveys. However, at present there is no compelling case for introducing the standard use of diaries in LSMS surveys.

#### *Data from Other Parts of the Questionnaire*

While a great deal of the data required to calculate consumption aggregates come from the consumption module of the household questionnaire, other important data are usually collected in other modules of the survey. These are reprised briefly here to serve as a checklist for the overall survey design. If survey designers decide not to collect data on the items mentioned here in the other modules, the designers should ensure that data on these items are collected in the consumption module.

**ROSTER.** Analysts need to know the number of household members in order to compute per capita expenditure measures. If analysts intend to calculate equivalence scales, they need data on the age and sex of the members as well. Data on age, sex, and number of household members are never omitted from rosters.

#### **HOUSING, WATER, SANITATION, AND FUEL MODULES.**

The housing module will collect most of the information needed to impute the use value of owner-occupied housing. It is also the usual place to gather information on utilities (such as electricity, piped gas, and telephone service) and on expenses for water, sanitation, and some kinds of fuel—although it is now being suggested that in some surveys these expenses may be moved to separate extensive water, sanitation, and fuel modules. (Such modules are discussed in Chapter 14.)

**EDUCATION MODULE.** Detailed data on household expenditures on school fees, uniforms, textbooks, supplies, bus fares, and so on are usually collected in the education module for each student, using recall periods that are considered to be appropriate for the category of expenditure. These expenditure data are used to compute consumption aggregates. Collecting this information in the education module makes it easier for the questions to refer to specific individuals, which is necessary for much of education analysis. The fact

that detailed information is collected also means that the information is more likely to be complete. Conceptually, there may be some overlap between some of the items recorded in the education and consumption modules. For example, does the category on children's clothing in the consumption module include or exclude school uniforms? Few LSMS questionnaires have been careful to specify this in the questionnaire itself, although it may be addressed during the interviewer training process.

**HEALTH MODULE.** Detailed data on health care expenditures such as payments to doctors or other medical professionals for prescription medicines and for lab tests are usually collected in the health module for each person who incurred such expenses during the recall period, which is usually the previous four weeks. Data on expenditures on health insurance and over-the-counter medicines are usually collected in the consumption module. Some surveys have also included a general question or two in the consumption module about expenditures on the items covered in more detail in the health module. In the consumption module the coverage is for the whole household and for a reference period of up to a year. The detailed data that are usually collected in the health module will generally give higher means than the aggregate questions in the consumption module, but the shorter recall period in the health module will result in higher variances. Analysts may then choose which method best suits their particular analysis. However, survey designers must ensure that at least one of the two data sets is included. If only one is to be included, the detail provided in the health module is probably preferable because it will yield more accurate means and support many health sector analyses.

**EMPLOYMENT MODULE.** The employment module is usually the place to gather information on households' consumption of goods provided in kind as a part of wages, as well as on household members' commuting expenses. Different analysts tend to handle commuting and other job-related expenses (such as child care, uniforms, and fees for professional associations) differently. Some even exclude these expenses altogether from the consumption total on the grounds that they do not increase the household's welfare. At any rate, it is useful to gather data on commuting expenses to give analysts the choice.

**ENVIRONMENT MODULE.** Households sometimes obtain important resources from the environment. Firewood and water for household use are perhaps the most familiar examples, although a wide range of plants, animals, and minerals can be gathered from wild or common property for use as food or fodder or as inputs into the household's enterprises, agriculture, or housing. To gather a full range of data on households' use of such resources would probably require a special purpose module and the modification of the consumption, agriculture, household enterprise, and housing modules to ensure full accounting while avoiding double-counting. This has yet to be done in an LSMS survey, though it has been done in some interesting single-purpose surveys (see Cavendish 1998). However, several past LSMS surveys have gathered information on at least a few of these resources, especially water and firewood. Such questions have typically covered the quantities used by the household, as well as in some cases the time spent collecting or carrying the goods or the distance traveled to find them. Survey designers need to decide how much emphasis to give to this issue in the survey; they also need to review all pertinent modules to ensure that modifications have been made where needed.

**COMMUNITY AND PRICE QUESTIONNAIRES.** The community and price questionnaires are the places to gather data on prices at the community level. If neither a community nor a price questionnaire is included in the survey, it is critical that either adequate regional price indices are available from some other source or—more likely—data on quantities of food and fuel consumed by households are obtained in the consumption module. It is also vital that the survey designers be willing to base regional price adjustments only on food, not using nonfood in the price index.

**SAVINGS MODULE.** Data on any regular saving by households can be gathered in either the savings module or the consumption module.

**INTERHOUSEHOLD TRANSFERS.** Many surveys have included a separate module to elicit data on flows of transfers in and out of the household. In general, it makes sense to put the questions about inflows and outflows of transfers together and to elicit parallel information about them. In this book, the inflows are discussed in the chapter on transfers and other nonla-

bor income (Chapter 11) and a page on outgoing transfers is included in the draft consumption module introduced in the following section of this chapter (and provided in Volume 3).

### **Draft Modules**

Volume 3 presents one version of a consumption module to be included in an LSMS-type household survey. The draft module finds a middle ground in terms of length. Some questions could conceivably be deleted, and other areas could be explored in much more detail. When and how to make such adjustments is explained in the fourth and final section of the chapter.

Survey designers need not organize the submodules in exactly the order in which they are presented here. There seems to be no rigorous evidence about how different systems of organization affect the answers given by respondents. However, there are a few principles of organization that survey designers should bear in mind, as discussed below.

#### *Organization*

There are different ways of grouping items together: by type of item (such as food, fuel, or clothing), by place of purchase, by recall period, or by the kind of follow-up questions asked (about quantities, prices, or home production). The exact layout will depend upon the circumstances in the country of the survey—including what people consume and how they acquire it—and on the objectives of the survey. The survey designers must draft one or more versions of a questionnaire and field-test them rigorously to determine whether the proposed module actually works.

#### *Comprehensiveness*

The questionnaire should cover all types of consumption. It is often pragmatic to define some categories that contain groups of items (for example, "canned goods") in order to produce a comprehensive list of expenditures without making the list unduly long.

#### *Specificity*

It is important to list certain items individually. As a rule, items should be listed individually if they are important sources of calories or are particularly interesting to analysts in their own right. Examples include subsidized fuels and goods whose consumption is par-

**Box 5.1 Cautionary Advice**

- *How much of the draft module is new and unproven?* The draft consumption module presented here closely follows the approach taken in many previous LSMS surveys.
- *How well has the module worked in the past?* The consumption modules in most previous LSMS surveys have produced data that have appeared to many analysts to have reasonable magnitudes and to show expected patterns of consumption by categories of items and over types of households. However, field tests of this module in some countries in Africa showed that respondents had difficulty reporting data on quantities of goods purchased, so questions about quantities purchased were deleted in those countries. In some surveys respondents have seemed to be confused about whether to report a total value and a quantity or a value per unit of purchase and a quantity. Where this confusion arose, it was necessary to delete some observations from the resulting data set or even to discard all of the data with a short recall period. To minimize this risk, the questions should be clearly formulated, the order of the questions should be thoroughly field tested, and the interviewers should be fully trained. However, the most effective insurance would be to use two recall periods, at least for food—thus ensuring that there is always a backup measure of consumption in the data set.
- *Which parts of the module most need to be customized?* Survey designers should base their decisions about which items to include in the different submodules on local consumption patterns. If a comparable survey was previously fielded in the country and if the consumption module in that survey was designed broadly in accordance with the best practices described in this chapter, the designers of the new survey should use the same recall periods and the same degree of item disaggregation as in the previous survey.

ticularly favored or disfavored by policymakers (“merit” goods and “sin” goods).

**Double-Counting**

It may sometimes be useful to create a cross-check by gathering data on the same items in more than one place in the questionnaire. In these cases, the questions should be worded carefully so that the analyst knows how to exclude one or the other to avoid double-counting these data in the estimate of total consumption. For example, if purchases of “major kitchen appliances” are listed among consumption expenditure

items, the inventory of durable goods should not list “stoves” since it would not be clear whether stoves were included among “major kitchen appliances.” In this case the durable goods inventory should either use the same category—“major kitchen appliances”—or omit all such items. This rule applies for cross-checks within the same module or across two modules.

**Customization**

Each questionnaire must be customized to reflect the circumstances in the country of the survey. The lists of consumption items must be customized to reflect local consumption patterns and local terminology. The recall periods for these items should depend on the shopping patterns of the local people.

**Annotations to the Questionnaire**

This section consists of a series of notes and annotations to each of the draft submodules introduced in the previous section and presented in Volume 3. These notes are followed by several general guidelines for dealing with some special circumstances within which a survey may operate. The section ends with some advice about how to keep the module as short as possible, consistent with the need to gather sufficient data to support the analysis of the most important policy issues in the country.

**Part A: Daily Expenditures**

The items included in this submodule will vary from country to country. The idea is to capture (using a short recall period) the small, repetitive miscellaneous transactions that many people engage in almost every day. Individuals often purchase the items listed here with their “walking around money.” As such, this is one of the places in the consumption module where it would be feasible to collect individual-specific data.

The list of items used in the draft submodule is a mixture of the lists from various previous LSMS surveys. Other items that could be included are: khat, flowers, gasoline, firewood, haircuts, shaves, baths, and tips.

In this submodule, there is a special grid for meals consumed away from home. Experience has shown that the more detail appears in the grid, the higher the numbers reported will be. In Jamaica the 1993 survey used questions that were individual-specific, and the estimate of expenditures on food purchased away from home

accounted for 5 percentage points more total food consumption than the 1992 survey, which used a single question on the daily expenditures grid (Table B-6 in Statistical Institute and Planning Institute of Jamaica 1994; Table B-4 in Statistical Institute and Planning Institute of Jamaica 1995). If the grid is individual-specific it is generally best for each person to respond for him or herself rather than to use a single respondent as specified in the draft submodule. Such an individual-specific page should be adjacent to other individual-specific modules. Care may also be needed to avoid double-reporting subsidized lunches in the factory canteen or the schoolyard, as these may also be captured in the employment or education modules.

This draft submodule is modeled after the one used in the Kazakhstan questionnaire because the submodule from Kazakhstan yielded mostly answers that made intuitive sense. The number of meals reported was mostly divisible by five, indicating regular patterns associated with the work week, and the unit value for the meals seemed plausible.

#### **Part B: Food and Fuel**

This submodule gathers information on the food and fuel consumed by the sample households.

**LIST OF ITEMS.** The list of items that should be included in this submodule will vary from country to country. The full-scale expenditure survey done every five or ten years in most countries to provide weights for the consumer price index has served as the basis for the development of the lists of foods in past LSMS surveys. Special care should be taken to itemize goods that contribute substantially to the total number of calories consumed and to expenditures, as well as goods that are (or are most likely to be) subsidized. Thus, in Central America, rice, beans, and tortillas should each be listed separately rather than in a group of “starches” or “basic grains.”

To prevent the list of items from becoming unwieldy (in the order of hundreds), there will have to be some grouping of items, using categories such as “canned foods” or “vegetables.” These groupings, while necessary, can pose problems for interviewers in eliciting answers about quantities of the item consumed and for analysts in trying to establish the nutrient content of the items. There is no perfect solution to this problem. For some items it may be pragmatic to black out the quantities question, thereby eliminat-

ing a source of confusion in the interview at the small cost of losing a piece of ambiguous data. For other items, taking care to keep the subgrouping reasonably homogenous may help resolve the issues of establishing quantities and estimating caloric content. “Leafy green vegetables” and “potatoes, sweet potatoes, and other tubers” are more internally homogenous categories than “vegetables.” Whether or not they adopt these approaches, survey designers need to bear the problems of grouping in mind.

It is usual, and probably helps respondents remember the necessary information, to list similar items together. In past surveys this has usually meant placing botanically similar items side by side—in other words, all meats together and all fruits and vegetables together—and, within each group, placing first the items that are more commonly consumed. Such placement probably reflects people’s shopping patterns in most countries; vegetables may be in one section of the market and meats in another, or people may buy their vegetables from the greengrocer and their meat from the butcher. If a very long list of items is used, it might make the interview more manageable to divide the list into subgroups in a more explicit way than is shown in the draft submodule, with subheadings for each subgroup. Then each subgroup could have a filter question such as “did your household purchase any meat since my last visit?” After that, separate questions on each kind of meat would follow. Some surveys supplement this general filter question by showing the respondent a card that either lists or illustrates (depending on the degree of literacy in the country) the various items in the subcategory as a prompt. However, no past LSMS survey has tried this.

The list of foods used in the draft submodule as an illustration came from the Pakistan LSMS questionnaire (with the addition of beer and other alcoholic beverages, which are important items of consumption in many countries, though not in Pakistan). Thus the list is specific to Pakistan and includes items (dal, gur, and ghee) that may be inappropriate to list in other countries. The list also omits some other items that should be included in other countries, such as pork, cassava, yams, tomatoes, papayas, and bananas. The key point is that the list must be customized to reflect local patterns of food consumption.

**RECALL PERIOD.** The recall period used in this submodule is the time since the interviewer’s last visit to

the household. Questions in the food submodule presume that the interviewer's last visit to the household was about two weeks before the interview on consumption; this was generally the case in the old prototypical fieldwork plan where each primary sampling unit was visited by the survey teams twice, two weeks apart, and different sections of the questionnaire were administered in each visit. The two-visit routine served a number of functions other than providing the recall period; however, now that data entry operators often travel with the field teams rather than remaining in a regional office, the logistical reasons for the two separate visits are being eroded. Moreover, in many surveys, especially ones that have shorter questionnaires than the full LSMS survey, only one visit is ever made to the primary sampling unit and the entire questionnaire is used in that visit. In single-visit scenarios the recall question (Question 2) should be changed from "since my last visit" to "in the past two weeks." In countries with high inflation rates or where other expenditure surveys use a shorter reference period, survey designers may shorten the recall period to a week or to whatever period matches that of the other survey.

**BARTER.** If barter is common in the country of the survey, it can be included in the wording of the purchase question (Question 5). If barter is very important, survey designers may wish to give it a question of its own. It will be particularly important to pilot-test the wording of the purchase question as this can often be awkward.

**QUANTITY.** The quantity question (Question 4) applies to the short recall period question, so that when a unit value is derived from this data, it refers to a specific time period. To have a unit value (quasi price) for some indeterminate month during the year is not very helpful where there is even modest inflation or seasonal variation in prices.

**ORDERING OF QUANTITY AND EXPENDITURE.** There is some debate about the proper ordering of these two questions (Questions 3 and 4). Essentially, the field test should be the guide. Expenditure has been placed first in the draft submodule for two reasons. First, it is the more important piece of data. Second, collecting expenditure first is likely to reduce the risk of ambiguity or misunderstanding on the part of the respondent.

If questions were in the reverse order, a respondent might give the interviewer an answer that referred to the household's expenditure per unit of quantity rather than for the total purchase. Instead of "I spent 50 pesos on meat. I bought two kilos," the informant might answer "I bought two kilos" and "I paid 25 pesos." The respondent would mean 25 pesos for each kilo, not in total, and the true response would be misrecorded. In Pakistan this problem occurred often enough to call into question the accuracy of all of the short period data. However, in the Panama field test, the questions seemed to work better with the quantity question placed before the expenditure question.

**FUEL.** Fuel is placed in the same submodule as food because it is the main category of nonfood that is purchased in convenient, standard units. This follows the principle that wherever quantities can be collected easily and accurately, they should be. Hence it is convenient, though not usual, to put fuel in the food submodule so that the follow-on questions on quantities will apply.

**HOME PRODUCTION.** There are several options about how to arrange the questions on home production (Questions 7–9). The option most commonly used in previous LSMS surveys has been to place these questions either together with food purchases, as is done here, or in a separate submodule on home production. Occasionally the consumption of home-produced food has appeared in the agricultural module as part of the means by which a crop is disposed of. Since agriculture modules are becoming increasingly infrequent in LSMS surveys and since they may not always include crop disposal questions, the following discussion concentrates on the first two options.

Whether the questions on households' consumption of home-produced food should be placed adjacent to the questions on their purchases of food or in a separate submodule is an issue that will be affected by both the percentage of households that produce their own food and the range of goods produced in this way. If questions on home-produced goods are placed in a separate submodule, it is simple to add the filter question, "Has your household consumed any food produced at home?" Then it is possible to avoid asking a long list of inapplicable questions to households that do not produce any food at home. This will be most appropriate when home production is not

very common. Ultimately, what works best should be determined by the field test. Do respondents find it easier to think about all of the sources of a single food-stuff together? (For example, I bought some tomatoes, I grew some tomatoes in my garden, or my coworker gave me some tomatoes?) Or is it easier for them to think about their food budget in terms of the source of that food? (For example, when I shop I buy beans, tortillas and milk, but in my garden I grow tomatoes, papaya, and bananas.) The decision about whether to put these questions in a separate submodule will not materially affect the length of the interview, although separating home production from purchased food may take up more paper and make the questionnaire physically longer.

It is not necessary to have a home production question for every food item in the list, since some are industrially manufactured items that cannot be made in people's homes. Even some items that are not industrially manufactured are not commonly produced in people's homes and can likewise be omitted from the list of home-produced items. In the draft of the submodule in Volume 3, the boxes are blacked out in the home-production columns for items that are unlikely to be produced at home.

The quantities of food produced at home should be regarded as the most important data gathered about home food production. The quantity of a good is a factual, observable piece of information. Questions about value are more hypothetical, since by definition a home-produced item does not pass through the market. Theoretically, a farmgate price would be too low and a market price too high. There seems to be no evidence as to which price the consumer of home-produced food is likely to know more often or which answer he or she would give in response to a general question on the "value of home production." The draft submodule includes a general (ill-defined) question on the "value of home production" on the grounds that—for goods that have markets—the food purchases questions, community questionnaire, or both gather information on prices or unit values from food purchases and may gather data on farmgate prices in the agricultural module (if there is one). Thus the "value of home production" contributes a different set of information, at least on the aspect that the respondent feels is most pertinent.

**GIFTS.** In some past LSMS surveys, the gifts section included the value of food received as payment in kind

from employers. This is not included in this draft submodule, because it already appears in the various employment submodules.

**USE OF ENVIRONMENTALLY PROVIDED GOODS.** If survey designers wish to try to put an explicit value on the use of all environmentally provided resources, the list of food items should include any food items that are likely to be gathered, fished, or hunted. In addition, the wording of the questions may be changed or additional questions may be added to clarify that environmentally provided items are included in the questioning.

### *Part C: Nonfood Consumption*

Again, the list of items to be included in this submodule is country-specific and can usually be derived from the survey used to weight the consumer price index. The list is likely to contain more groups of items (for example, "clothing") than were contained in the food list. The nonfood list should be designed with the following objectives: to cover all aspects of the household's budget that are not covered elsewhere in the household questionnaire; to cover these in a logical manner that respondents find congenial; to gather information on specific items that may receive heavy subsidies or attract heavy taxes (for example, kerosene, gasoline, tobacco, or alcohol); and, sometimes, to gather the data in a manner that will enable analysts to study intrahousehold issues.

For most items it will not be practical to collect data on quantity (or, implicitly, unit values) due to the difficulties of establishing meaningful units, especially for groups of goods. For items where meaningful units can be established (usually fuel), data on the quantity consumed should be gathered.

Note that the boundary between consumption items and durable goods is somewhat fuzzy. Many of the items usually listed in the consumption module (and included in this draft submodule) may actually last for more than a year. Kitchen equipment (including cups, forks, plates, and saucepans), furniture (including beds, tables, cupboards, chairs, and rugs), and linens (including sheets, towels, and blankets) all appeared in the consumption module of the Côte d'Ivoire questionnaire rather than in the durable goods module, even though all of these items were likely to last for more than a year. Nonetheless, so many different items may be listed that it seems pragmatic to measure their consumption via the flow into the

household rather than trying to enumerate the whole stock item by item and then compute a use value for it.

The list of items in the draft submodule is derived from several different questionnaires from previous LSMS surveys, principally the surveys in Jamaica and Nepal.

The same considerations that apply for food apply to the barter and gift of nonfood goods and to nonfood goods (such as firewood) gathered from the environment. Home production is theoretically of interest but has rarely been specifically enumerated in past LSMS surveys. If survey designers decide that home-produced nonfood goods are important, they can be handled in a manner similar to that for home-produced food.

#### *Part D: Expenditures on Private Interhousehold Transfers*

Respondents may find it most logical to be asked questions about their income from and expenditures on interhousehold transfers in the same place in the questionnaire. In this book, however, questions on household expenditures on private interhousehold transfers are included here and questions on income from interhousehold transfers are included in the module on transfers and other nonlabor income (introduced by Chapter 11). The two submodules were developed together and can be positioned adjacent to one another in the household questionnaire in a separate module.

Two versions of this submodule are included. The short version is designed to obtain basic information when the consumption module is being shortened as much as possible and the study of private safety nets is not deemed important. The longer submodule includes questions about the recipients and their relationship to the donor or head of the household, as well as on the amounts, regularity, and purpose of the transfers. The answers to these questions supply analysts with a lot of information for studying transfers.

For Question 12 it would be useful to get as much detail on the destination as can easily be coded—for example, to the district or county level. For Questions 12 and 13 the codes should coincide with those used in the migration module. Questions 6 and 7 use the same codes as are used on the roster, which is why code 1 is missing. (On the roster module it is for the head of household.)

#### *Part E: Durable Goods*

The items in the durable goods list should be items that last substantially longer than a year and are so large in relation to the household's standard of living that they can be separately enumerated and respondents can accurately remember information about their purchase after several years have gone by. A car would meet this definition of a durable good but a shirt would not—even though both may last for several years.

The durable goods page in this submodule is divided into two blocks so that households can report on the value of two of the same kind of item (for example, two bicycles). Most past LSMS questionnaires have included a list of a dozen or two different kinds of durable goods, though the lists have usually been longer in the surveys fielded in the countries of the former Soviet Union. The most appropriate items to include in the durable goods list will vary from country to country. What is considered expensive enough to be a durable good will also vary from country to country. For example, cooking pots were on the durable goods list in the survey in Kagera, Tanzania, whereas the list in the Jamaican survey included satellite dishes. The most common durable goods may also vary by climate. In tropical countries air conditioners or fans will often be on the list, whereas in cold climates various sorts of heaters will be listed. Culture also plays an important role. Jewelry, carpets, and guns were listed in Pakistan but are unlikely to be listed in the durable goods submodule in many other countries.

#### *Special Circumstances*

This subsection provides survey designers with advice on how to deal with various special circumstances that they may face in planning and designing a survey in their particular country.

**DEALING WITH INFLATION.** In countries with high inflation it may be necessary to modify the questionnaire. (However, if inflation is very high, it is not clear how well these suggested modifications will work.)

First, the recall period should be shortened to the shortest period that is reasonable for the type of item in question. The recall period on food might change from the previous two weeks to the previous week. The “usual-month” questions on food might be dropped altogether, since it would be unclear to which time period they referred, making it impossible to

deflate the expenditures appropriately. The recall period for nonfood items might be shortened from a year to three or six months. Shortening the recall periods could increase the variance of the estimates, but in any case it would be impossible to interpret any means for data collected when prices were very different. Moreover, inflation will cause people to make more frequent purchases (so that the real value of their money does not diminish), so the tradeoff between biases in the mean and variance may be less than in places where inflation is low.

The second modification that can be made to take inflation into account involves asking respondents about the dates when they made certain large purchases so that analysts can deflate these figures appropriately. Third, if a currency other than the national currency has become a *de facto* unit of account, survey designers might allow respondents to give their answers in either the local or the international currency.

**CREDIT.** If the use of consumer credit is of particular analytical and policy interest in the survey country, a short module on households' use of credit (particularly for purchasing food) could be inserted after the food submodule, and questions on purchases on credit could be added to the durable goods submodule, as is suggested in Chapter 21 on credit.

**SUPPORTING INTRAHOUSEHOLD ANALYSIS.** If supporting intrahousehold analysis is a special goal of the survey, the consumption module can be reorganized so that at least some of the information is individual-specific. The main principle is to classify all items according to two criteria. The first criterion is whether they are consumed individually (like a taxi fare or a shirt) or jointly (like washing powder or television use). The second criterion is whether it is easy to distinguish if the item was consumed by one household member or another. Food eaten at home is individually consumed, but, since it is usually purchased and prepared for the household as a whole, it is difficult to distinguish how much is consumed by each individual. On the other hand, it is easier to assign the consumption of food consumed in restaurants or food stalls to particular individuals. After the items that are individually consumed and easily distinguished as being consumed by a given individual have been determined, these items can be listed on a separate grid that the interviewer fills out for each individual.

### *Keeping the Module as Short as Possible*

It is important to avoid tiring or annoying respondents with an overly long interview. However, there are relatively few options for shortening the consumption module, aside from using the shorter of the two versions of the submodule on interhousehold transfers. As explained in the second section of this chapter, reducing the number of items or categories for which information is collected can easily lead to an underestimation of consumption. Also, getting a comprehensive measure of consumption requires inquiring about purchases, home production, and gifts of all commodities consumed by the household, so it is not easy to reduce the number of questions about each item.

There are a few questions that could be cut in the draft submodules, though this would result in a loss of important data in each case. Probably the most expendable question is Question 7 of Part E, the last of the series for durable goods. Removal of this question would mean that the valuation of durable goods would rest solely on the assumption that the average life of the good is twice the average age reported across all households.

It would be possible to drop Question 9 in the food submodule, and instead to value home production using data from the expenditure questions as well as the community questionnaire. There is an element of risk in this, since households would no longer have the option to express their answers in value terms, which they may have found easier to use than quantity terms. Of course, this option is only available if there are questions on quantities in the food expenditure submodule of consumption or if data on prices are gathered in the community questionnaire for all items on the home-produced list. Alternatively, survey designers could drop the quantity questions about purchases and home production. (Note that the quantity and the value questions on home production cannot both be dropped simultaneously.) However, quantity is an important piece of data in its own right, so dropping quantity questions reduces analytical potential. Most importantly it means that prices from some other sources are required to derive a calorie-based poverty line from the data set.

A final option would be to drop the dual recall periods in the food submodule or the nonfood submodule. Doing this would be risky, since if something went wrong in the fieldwork analysts would be unable

to use the data set to calculate what is arguably the most important variable in the whole enterprise—the total household consumption. Because food is more important in the total than nonfood, it is preferable to drop the dual recall periods for nonfood items before dropping those for food items.

## Note

The authors are grateful for useful comments to Richard Blundell, Martin Browning, Paul Glewwe, John Hoddinott, Alberto Martini, Andrew McKay, Raylynn Oliver, and, especially, Christopher Scott. The Statistical Institutes of Armenia, Ecuador, El Salvador, Jamaica, and Latvia, and the Kiev Institute of Sociology in Ukraine conducted experiments in the collection of consumption data that are summarized. Eric Edmonds, Dean Jolliffe, Kinnon Scott, Diane Steele, Tilahun Temesgen, and Wlodek Okrasa contributed to the analysis that underlies some of the discussion.

## References

- Ahmad, S. Ehtisham, and Nicholas H. Stern. 1991. *The Theory and Practice of Tax Reform in Developing Countries*. Cambridge: Cambridge University Press.
- Atkinson, Anthony B., and John Mickelwright. 1983. "On the Reliability of Income Data in the Family Expenditure Survey 1970–77." *Journal of the Royal Statistical Society Series A* 146: 33–61.
- Baker, Judy. 1996. Personal correspondence. Washington, D.C.
- Behrman, Jere R., and Anil B. Deolalikar. 1987. "Will Developing Country Nutrition Improve with Income? A Case Study for Rural South India." *Journal of Political Economy* 95 (June): 492–507.
- Bhalla, Surjit S. 1979. "Measurement Errors and the Permanent Income Hypothesis: Evidence from Rural India." *American Economic Review* 69: 295–307.
- . 1980. "The Measurement of Permanent Income and its Application to Saving Behavior." *Journal of Political Economy* 88: 722–43.
- Bhattacharyya, Nikhilesh. 1963. "On the Effects of Itemization in the Family Budget Schedule." Indian Statistical Institute, Calcutta.
- Biemer, Paul P., Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, and Seymour Sudman, eds. 1991. *Measurement Errors in Surveys*. New York: Wiley.
- Blaizeau, Didier. 1998. "Household Expenditure Surveys in the Seven UEMOA Countries." Institut National de la Statistique et des Etudes Economiques, Paris.
- Bouis, Howarth E. 1994. "The Effect of Income on Demand for Food in Poor Countries: Are Our Databases Giving us Reliable Estimates?" *Journal of Development Economics* 44 (June): 199–226.
- Bouis, Howarth E., and Lawrence J. Haddad. 1992. "Are Estimates of Calorie–Income Elasticities too High? A Recalibration of the Plausible Range" *Journal of Development Economics* 39: 333–64.
- Bourguignon, François, and Pierre-André Chiappori. 1992. "Collective Models of Household Behavior: An Introduction." *European Economic Review* 36 (April): 355–64.
- Bourguignon, François, Martin J. Browning, Pierre-André Chiappori, and Valérie Lechene. 1993. "Intrahousehold Allocation of Consumption: A Model and Some Evidence from French Data." *Annales d'Économie et de Statistique* 29 (January–March): 137–56.
- Bradburn, Norman M., Janellen Huttenlocher, and Larry Hedges. 1994. "Telescoping and Temporal Memory." In Norbert Schwartz and Seymour Sudman, eds., *Autobiographical Memory and the Validity of Retrospective Reports*. New York: Springer-Verlag.
- Branch, E. Raphael. 1994. "The Consumer Expenditure Survey: A Comparative Analysis." *Monthly Labor Review* 117 (12): 47–55.
- Browning, Martin, François Bourguignon, Pierre-André Chiappori, and Valérie Lechene. 1994. "Income and Outcomes: A Structural Model of Intrahousehold Allocation." *Journal of Political Economy* 102 (December): 1067–96.
- Cavendish, William. 1998. "The Complexity of the Commons: Environmental Resource Demands in Rural Zimbabwe." University of Oxford, Centre for the Study of African Economies, Oxford.
- Coder, John. 1991. "Exploring Non-Sampling Errors in the Wage and Salary Income Data from the March Current Population Survey." Bureau of the Census, Household and Household Economic Statistics Division, Washington, D.C.
- Cramer, Jan S. 1969. *Empirical Econometrics*. Amsterdam: North-Holland.
- Dandekar, V. M., and N. Rath. 1971a. *Poverty in India*. Pune: Indian School of Political Economy.
- . 1971b. "Poverty in India: Dimensions and Trends." *Economic and Political Weekly* 6: 25–48.
- Deaton, Angus. 1988. "Rice Prices and Income Distribution in Thailand: A Non-Parametric Analysis." *Economic Journal* 99 (Supplement): 1–37.
- . 1997. *The Analysis of Household Surveys: a Microeconomic Approach to Development Policy*. Baltimore, Md.: Johns Hopkins University Press.
- Deaton, Angus S., and John Muellbauer. 1980. *Economics and Consumer Behavior*. Cambridge: Cambridge University Press.
- Deaton, Angus S., and Christina H. Paxson. 1998. "Economies of Scale, Household Size, and the Demand for Food." *Journal of Political Economy* 106 (October): 897–930.
- Deaton, Angus S., and Salman Zaidi. 1999. "Guidelines for Constructing Consumption Aggregates for Welfare Analysis."

- Princeton University and Development Research Group, World Bank, Washington, D.C.
- Eisenhower, Donna, Nancy A. Mathiowetz, and David Morganstein. 1991. "Recall Error: Sources and Bias Reduction Techniques." In Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, and Seymour Sudman, eds., *Measurement Errors in Surveys*. New York: Wiley.
- Foley, M. 1996. Personal correspondence.
- Ghose, Suchismita, and Nikhilesh Bhattacharyya. 1993. "Engel Elasticities of Clothing and Other Items." *Sarvekshana* 17: 35–9.
- . 1995. "Effect of Reference Period on Engel Elasticities of Clothing and Other Items: Further Results." *Sankhya: The Indian Journal of Statistics* 57 (Series B): 433–49.
- Gibson, John. 1998. "How Robust are Poverty Comparisons to Changes in Household Survey Methods? A Test Using Papua New Guinean Data." University of Waikato, Department of Economics, New Zealand.
- Gieseman, Raymond. 1987. "The Consumer Expenditure Survey: Quality Control by Comparative Analysis." *Monthly Labor Review* 110 (March): 8–14.
- Glewwe, Paul. 1987. *The Distribution of Welfare in Peru in 1985–86*. Living Standards Measurement Study Working Paper 42. Washington, D.C.: World Bank.
- Glewwe, Paul, and K.A. Twum-Baah. 1991. *The Distribution of Welfare in Ghana, 1987–88*. Living Standards Measurement Study Working Paper 75. Washington, D.C.: World Bank.
- Griliches, Zvi, and Jerry A. Hausman. 1986. "Errors in Variables in Panel Data." *Journal of Econometrics* 31 (February): 93–118.
- Grootaert, Christiaan. 1986. "The Use of Multiple Diaries in a Household Expenditure Survey in Hong Kong." *Journal of the American Statistical Association* 81: 938–44.
- . 1993. "The Evolution of Welfare and Poverty Under Structural Change and Economic Recession in Cote d'Ivoire, 1985–88." Policy Research Working Paper 1078. World Bank, Washington, D.C.
- Grosh, Margaret. 1997. "The Policymaking Uses of Multi-topic Household Survey Data: A Primer." *World Bank Research Observer* 12 (August): 137–60.
- Grosh, Margaret, and Juan Muñoz. 1996. *A Manual for Planning and Implementing the Living Standards Measurement Study Surveys*. Living Standard Measurement Study Working Paper 126. Washington, D.C.: World Bank.
- India, Planning Commission. 1993. "Report of the Expert Group on Estimation of Proportion and Number of Poor." Perspective Planning Division, New Delhi.
- IMF (International Monetary Fund). 1995. *International Financial Statistics Yearbook*. Washington, D.C.
- . 1996. *International Financial Statistics Yearbook*. Washington, D.C.
- Jacobs, Eva, and Stephanie Shipp. 1993. "A History of the US Consumer Expenditure Survey: 1935–36 to 1988–89." *Journal of Economic and Social Measurement* 19: 59–96.
- Jolliffe, Dean, and Kinnon Scott. 1995. "The Sensitivity of Measures of Household Consumption to Survey Design: Results from an Experiment in El Salvador." World Bank, Policy Research Department, Washington, D.C.
- Lanjouw, J., and P. Lanjouw. 1996. "Aggregation Consistent Poverty Comparison: Theory and Illustrations." Yale University and World Bank, New Haven, Conn., and Washington, D.C.
- Laraki, Karim. 1989. "Ending Food Subsidies: Nutritional, Welfare, and Budgetary Effects." *World Bank Economic Review* 3 (September): 395–408.
- Mahalanobis, P.C. 1946. "Recent Experiments in Statistical Sampling in the Indian Statistical Institute." *Journal of the Royal Statistical Society* 109: 325–78.
- Mahalanobis, P. C., and S. B. Sen. 1954. "On Some Aspects of the Indian National Sample Survey." *Bulletin of the International Statistical Institute* 34.
- Martini, Alberto, and Anna Ivanova. 1996. "The Design and Implementation of the Income and Expenditure Survey of Belarus." Urban Institute, Washington, D.C.
- McWhinney, Isabel, and Harold E. Champion. 1974. "The Canadian Experience with Recall and Diary Methods in Consumer Expenditure Surveys." *Annals of Economic and Social Measurement* 3/2: 411–35.
- Minhas, B. S. 1988. "Validation of Large-Scale Sample Survey Data: Case of NSS Household Consumption Expenditure." *Sankhya* Series B, 50 (Supplement).
- Minhas, B. S., and S. M. Kansal. 1989. "Comparison of NSS and CSO Estimates of Private Consumption: Some Observations based on 1983 Data." *Journal of Income and Wealth* 11 (January): 7–24.
- Musgrove, Philip. 1978. *Consumer Behavior in Latin America*. Washington, D.C.: Brookings Institution.
- . 1979. "Permanent Household Income and Consumption in Urban South America." *American Economic Review* 69: 355–68.
- Neter, John. 1970. "Measurement Errors in Reports of Consumer Expenditures." *Journal of Marketing Research* 7: 11–25.
- Neter, John, and Joseph Waksburg. 1964. "A Study of Response Errors in Expenditure Data from Household Interviews." *Journal of the American Statistical Association* 59 (305): 18–55.
- Newbery, David M. G., and Nicholas H. Stern, eds. 1987. *The Theory of Taxation for Developing Countries*. New York: Oxford University Press.
- Paxson, Christina H. 1992. "Using Weather Variability to Estimate the Response of Savings to Transitory Income in Thailand." *American Economic Review* 82: 15–33.

- . 1993. "Consumption and Income Seasonality in Thailand." *Journal of Political Economy* 101: 39–72.
- Reagan, Barbara B. 1954. "Condensed versus Detailed Schedule for Collection of Family Expenditure Data." U.S. Department of Agriculture, Agricultural Research Service, Washington, D.C.
- Rubin, David C., and Alan D. Baddeley. 1989. "Telescoping is not Time Compression: a Model of the Dating of Autobiographical Events." *Memory and Cognition* 17: 653–61.
- Sahn, David. 1989. *Seasonal Variation in Third World Agriculture: the Consequences for Food Security*. International Food Policy Research Institute. Baltimore, Md.: Johns Hopkins University Press.
- Scott, Christopher, and Ben Amenuvege. 1990. "Effect of Recall Duration on Reporting of Household Expenditures: An Experimental Study in Ghana." Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper 6. World Bank, Washington, D.C.
- Scott, Christopher, Mary Strode, and Oliver Chinganya. 1998. "Estimation of Annual Expenditure from Monthly Survey Data."
- Scott, Kinnon. 1994. "Venezuela: Poverty Measurement with Multiple Data Sets." World Bank, Development Research Group, Washington, D.C.
- . 1997. Personal conversation. Washington, D.C.
- Scott, Kinnon, and Wlodzimierz Okrasa. 1998. "Analysis of Latvia Diary Experiment." World Bank, Development Research Group, Washington, D.C.
- Sen, Amartya K. 1999. *Development as Freedom*. New York: Knopf.
- Silberstein, Adriana R. 1990. "First-Wave Effects in the US Consumer Expenditure Interview Survey." *Survey Methodology* 16: 293–304.
- Silberstein, Adriana R., and Stuart Scott. 1991. "Expenditure Diary Surveys and their Associated Errors." In Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, and Seymour Sudman, eds., *Measurement Errors in Surveys*. New York: Wiley.
- . 1992. "Seasonal Effects in the Reporting of Consumer Expenditure." *Proceedings of the American Statistical Association Survey Research Section* 333–8.
- Srinivasan, T. N. 1994. "Database for Development Analysis: An Overview." *Journal of Development Economics* 44 (June): 3–27.
- Statistical Institute and Planning Institute of Jamaica. 1994. *Jamaica Survey of Living Conditions 1992*. Kingston.
- . 1995. *Jamaica Survey of Living Conditions 1993*. Kingston.
- . 1996. *Jamaica Survey of Living Conditions 1994*. Kingston.
- Steele, Diane, 1998. "Ecuador Consumption Items." World Bank, Development Research Group, Washington, D.C.
- Strauss, John, and Duncan Thomas. 1995. "Human Resources: Empirical Modeling of Household and Family Decisions." In Jere Behrman and T. N. Srinivasan, eds., *Handbook of Development Economics*. Amsterdam: Elsevier.
- Subramanian, Shankar, and Angus S. Deaton. 1996. "The Demand for Food and Calories." *Journal of Political Economy* 104 (February): 133–62.
- Sudman, Seymour, and Robert Ferber. 1971. "Experiments in Obtaining Consumer Expenditures by Diary Methods." *Journal of the American Statistical Association* 66: 725–35.
- Sudman, Seymour, Norman N. Bradburn, and Norbert Schwarz. 1996. *Thinking About Answers*. San Francisco: Jossey-Bass.
- Swindale, Anne, and Punam Ohri-Vachaspati. 1997. "Household Food Consumption Indicator Guide." International Science and Technology Institute, IMPACT Food Security and Nutrition Monitoring Project, Washington, D.C.
- Tsoflias, Lynn. 1996. Personal communication. Washington, D.C.
- Tucker, Clyde. 1992. "The Estimation of Instrument Effects on Data Quality in the Consumer Expenditure Survey." *Journal of Official Statistics* 8: 41–61.
- Tucker, Clyde, and Claudette Bennett. 1988. "Procedural Effects in the Collection of Consumer Expenditure Information." *Proceedings of the Section on Survey Research Methods*. American Statistical Association. 256–61.
- van der Gaag, Jacques. 1994. Personal communication, Washington, D.C.
- Webb, R., and G. Fernandez Baca. 1993. *Peru en Numeros 1993*. Anuario Estadístico. Cuanto. Lima, Peru.
- Wolpin, Kenneth I. 1982. "A New Test of the Permanent Income Hypothesis: The Impact of Weather on the Income and Consumption of Farm Households in India." *International Economic Review* 23 (October): 583–94.
- World Bank. 1993. "Indonesia: Public Expenditures, Prices, and the Poor." Report 11293-IND, Indonesia Resident Mission, Jakarta.
- . 1994a. "Guyana: Strategies for Reducing Poverty." Report 12861-GUA, Country Operations Division 2, Country Department 3, Latin America and the Caribbean Region. Washington, D.C.
- . 1994b. "Kingdom of Morocco: Poverty, Adjustment & Growth." Vol. 2. Report 11918-MOR. Middle East and North Africa Region, Country Department 1, Country Operations Division, Washington, D.C.
- . 1995a. "The Kyrgyz Republic. Poverty Assessment and Strategy." Report 14380-KG. Human Resources Division, Europe and Central Asia Country Department 3, Washington, D.C.
- . 1995b. "Republic of Nicaragua. Poverty Assessment." Volume 2. Report 14038-NI. Country Operations Division, Country Department 2, Latin America and the Caribbean Regional Office, Washington, D.C.

- . 1995c. "Tanzania Social Sector Review." Report 14039-TA. Poverty and Human Resources Division, Country Department 2, Africa Region, Washington, D.C.
- . 1996. "Jamaica Survey of Living Conditions (JSL) 1988-94 Basic Information." Poverty and Human Resources Division, Policy Research Department, Washington, D.C.
- . 1997. "Romania, Poverty and Social Policy." Report 16467. Human Resources Sector Operations Division, Country Department 2, Europe and Central Asia Region, Washington, D.C.
- Zambia, Central Statistical Office. 1995. *Household Budget Survey 1993/94*. Lusaka.



# 6 Household Roster

*Paul Glewwe*

Almost every household survey instructs interviewers to make a list of all the members of each household in the survey. This list is often called the household roster. In LSMS and similar multi-topic surveys, the household roster serves three distinct functions. It determines who is and who is not a member of the household. It collects some basic information on each member of the household. And it can be used to collect information on close relatives of household members—parents, children, spouses, siblings—who are no longer, or have never been, members of the household. The collection of information on these household “associates” is unusual, and thus distinguishes LSMS-type surveys from most other household surveys.

This chapter provides advice on how to design the household roster of an LSMS-type household survey. The first section discusses the three basic functions of the household roster. The second section introduces the draft household roster module (which is provided in Volume 3 of this book). The third section presents detailed explanatory notes on the draft module.

## **The Three Main Functions of the Household Roster**

The household roster has three basic functions in LSMS and other multitopic household surveys: determining household membership, collecting basic information on household members, and collecting information on nonresident family members.

### *Determining Household Membership*

Perhaps the most important function of the household roster is to determine which individuals are members of the household. This function is important because

the vast majority of information collected in the other modules of the household questionnaire pertains only to household members. If a household member is mistakenly classified as not belonging to the household, no information of any kind will be gathered about that person in the rest of the household questionnaire.

The first step in determining household membership is to define what is meant by a household. For the purposes of conducting a household survey, the standard definition of a household is a group of people who live together, pool their money, and eat at least one meal together each day (United Nations 1989).<sup>1</sup> While most people who live together do pool their money and eat together, there are some exceptions. For example, unrelated individuals may share a dwelling to minimize housing costs but may eat separately and not pool their money. In this case each individual should be considered as living in a separate household. A second example is when two distinct families live in the same dwelling but allocate some rooms to one family and the rest to the other. They

may or may not eat together, but if they do not pool their income they should be considered two separate households. A third example is household servants. Servants are generally not considered household members because they keep their income separate from the income of the household that employs them, even though they may eat some meals with the family. A final case is individuals who rent a room in a dwelling that belongs to one family. A renter may eat some meals with the family, but as long as the renter keeps his or her income separate from that of other household members, he or she should be considered a separate household.

The definition given above of what constitutes a household is clear for the vast majority of living situations. However, unusual situations arise in many countries. For example, in some polygamous societies a man may have several wives, each of whom has a separate dwelling. In this case each wife (and her children) constitutes a separate household. Which household the man belongs to is difficult to say. In this and other unusual cases flexibility is needed; one course of action is to adopt the definition of a household used in previous surveys in the particular country. An important principle to follow in these situations is that each person in the population should be assigned to one, and only one, household.

After a definition of what constitutes a household is settled, the next step is to assign each person in the population to one, and only one, household. If some people are not assigned to any household, they will not be represented in the survey; if other people are assigned to two or more households, they will be over-represented in the survey. The main difficulty is that over long periods of time, many people move from one household to another. The longer the period of time, the more common this is. To avoid these potentially troublesome situations, survey designers may wish to define household membership over a relatively short period of time. At one extreme, this might be the previous 24 hours—in effect, those who spent the night at the household surveyed—which would make the number of ambiguous cases quite small. On the other hand, many socioeconomic phenomena of interest to policymakers involve activities and situations that last for a considerable period of time. For example, several modules described in other chapters of this book, including the modules on consumption, income, employment, and agriculture, use a 12-month

recall period. To be consistent with these modules, it is recommended that LSMS and similar multitopic surveys define household membership over the previous 12 months.

How long must a person reside in a given household in order to be considered a member of that household? The main criterion for determining whether someone has been a household member during the previous 12 months is the number of months the individual has lived in the household during that time. A person who has lived in the household for all of the past 12 months should certainly be considered a household member. A person who has lived in the household for more than six months should also be considered a household member, since it is not possible for that person to be a member of any other household for more than six months. One potential problem is people who have lived in one household for exactly six months and in another household for six months. If household membership includes people who have lived in a household for 6 of the past 12 months, such persons have, in theory, a double chance of being surveyed. In practice this problem is probably minor because it applies to only a very small percentage of the population.

A more troublesome case is an individual who has lived in three different households during the previous 12 months—spending 4 months in each household. If household membership is defined as living in a household for 6 or more of the past 12 months, such individuals will never be sampled in the survey. However, if the threshold is 4 months or less, these people will be three times as likely to be included in the survey as people who lived in only one household during the previous 12 months. The prudent thing to do here is to err on the side of caution by setting a relatively low number of months as the threshold for considering a person to be a household member. The threshold set in many past LSMS surveys has been three or more months. This is prudent because as long as the number of months each person lived in the household during the previous 12 months is recorded in the household roster, analysts can decide for themselves whether to include as household members people who have spent 6 or fewer months in the household.

A related problem arises when people move between two or more households on a regular basis. For example, the main income-earner of a given household may work away from home during the

week but return home on weekends. In general, the rule to apply is to calculate the percentage of time that the person has been in the household during the previous 12 months, express this in terms of a number of months, and then apply the rule described above. One exception to this rule is a person whose time away from the household was clearly spent in an institutional setting, such as a workers' dormitory. Since that person is not in danger of being double-counted (because he or she did not live in any other household during the previous 12 months), he or she can be counted as a household member.

There have been some exceptions to the rule commonly used in past LSMS surveys that any person who has lived in the household for 3 or more of the past 12 months is considered to be a household member. First, a newly born infant is typically considered to be a household member even if he or she is less than three months old because the infant was obviously not a part of any other household during the other months. Second, any household member who has died during the previous 12 months is generally not considered a household member. The reason for this is that this person cannot provide answers for himself or herself, and it is often uncomfortable for other household members to answer extensive questions about someone in the household who has recently died. (An exception is that the fertility module introduced by Chapter 15 does collect information on any children who have died.)

Another possible exception is someone who has recently become a member of the household and clearly has not been a member of another household during the previous 12 months because he or she was living in some kind of institutional setting (for example, student housing, a military barracks, or a prison) or in a foreign country. For example, in the 1992–93 Vietnam LSMS survey, all recently demobilized soldiers were counted as household members even when they had only been members for one or two months. Another exception to the rule is that the person designated as the head of the household is always considered a household member regardless of the amount of time that he or she has spent in the household during the previous 12 months. A final exception that has been used in some previous LSMS surveys is that a new spouse, usually a wife, who has joined a household is a member even if he or she joined only one or two months ago. In Vietnam this rule was used because Vietnamese culture dictates that this person is a house-

hold member. Classifying these people as nonmembers may offend some members of the household.

#### *Collecting Basic Information on Household Members*

Once all of the household members have been identified, what information should be collected about them? The information should be limited to the most fundamental characteristics of each individual; data that pertain to other modules should be collected in those modules. The standard information to collect in the household roster is:

- The person's name.
- The person's date of birth (if known).
- The person's age (in years for adults, in months for children age 12 and under)
- The person's sex.
- The person's relationship to the head of household.
- The person's marital status.
- The number of months that the person has lived in the household during the previous 12 months.

This information is straightforward and uncontroversial. However, a few comments are required. The main purpose for obtaining the name of the person is to allow the interview to proceed smoothly and naturally. There is no reason to provide the names of household members to analysts who want to use the data.<sup>2</sup> This would violate the statistical regulations governing almost all official household surveys, which typically guarantee the confidentiality of information provided by respondents.

The reason for collecting data on the respondent's date of birth is that many people in developing countries, especially elderly people, have some difficulty remembering their exact ages. If a person's birth date is obtained, that information, along with the date of the interview, enables analysts to calculate the person's exact age (see Chapter 4 on metadata). Another reason for collecting both age and birth date information is that occasionally a person may unintentionally make an error when answering one of these questions. This information is sufficiently important that the responses to these two questions should be compared to check the accuracy of the ages of all household members.

It is important to obtain a record of the number of months that the person has lived in the household in question during the previous 12 months in order to ascertain who is, and who is not, a household member.

Many households include married couples, and it is useful in analysis to match each household member

to his or her spouse. In some households this will be obvious, but in households that contain extended families it may not be clear. Therefore, in general, each person who is married should be matched to the household member who is his or her spouse. If the spouse is not a household member, this can also be indicated in the household roster.

It is sometimes useful to collect other information about household members in the household roster, such as information on ethnicity, religion, and nationality. In countries where intermarriage between members of different ethnic groups is rare, questions about ethnicity need be asked only once, for the household as a whole. In this case the information should not be collected in the household roster but in the household identification page (see Chapter 4). If intermarriage between members of different ethnic groups is common, each person should be asked about ethnicity separately in the household roster—and a “mixed” category may be needed for the children of interethnic marriages. The same is true for religious affiliation; if it is rare to find people of different religions in the same household, this can be asked once for the whole household, but if it is common, each person should be asked individually in the household roster. Finally, in some countries a substantial fraction of the population may have migrated in from another country and thus consist of people who are not citizens. Data about citizenship can be collected in the household roster if noncitizens constitute a sizable fraction of the population.

Before moving to the next topic, it is useful to consider two other types of information that some survey designers may want to collect in the household roster: languages spoken and the recent location of household members. In some countries many people speak a language other than the official language of the country. Information on languages spoken may be particularly relevant for analyses of education and employment. One or two questions could be added to the household roster to record the languages spoken by each household member above a certain age. However, in most cases it is more convenient to collect this information in the education module. In particular, the standard education module in Chapter 7 asks household members to read a short sentence, and the notes to that module suggest that this could be done for more than one language. This is a natural place to ask a couple of questions about the languages spoken by household members.

Another piece of information that may be useful in analysis is whether each respondent is currently living in the household—more specifically, the number of days he or she has lived in the household during the past week or month. This would be useful for explaining why a particular individual could not answer questions for himself or herself, but it is even more useful for analyses of migration. In most surveys such questions should be included in the migration module.

#### *Collecting Information on Nonresident Family Members*

Most LSMS surveys collect information on the parents and children of household members regardless of whether these people are household members. This information is collected because it gives analysts and policymakers a better understanding of parent-child relationships within the household, because it illustrates the links between the household surveyed and other households, and because the data can be useful for applying certain econometric techniques, particularly the instrumental variables method.

Consider first the case of parents. As with spouses, when extended families live in the same household it is not always clear which children are associated with which adults. Thus it is useful to ask each person explicitly whether their mother and father are household members and, if so, which household members they are. If a person's mother or father is not a household member, the roster should gather some basic information on that individual—in particular, whether the individual is still alive, the individual's highest level of education, and the individual's main occupation. Survey designers may also want to ask respondents where their parents live if they are still alive. Such information is useful because parents' education levels are strongly associated with their children's levels of education, because the occupations of (adult) children are often correlated with their parents' occupations, and because many households depend on relatives to help them in times of need. (Thus a household's vulnerability to economic hardship may depend on whether household members' parents are alive or dead, as well as on where they live.)

Another issue is how to treat stepchildren and adopted children. Most stepchildren have one parent in the household and another who is dead or lives elsewhere. This situation will become clear from the information given on the child's parents, and thus needs no further discussion. Whether the parents of adopted children should be treated as their real parents depends

on the incidence and nature of adoption in the country surveyed. If adoption is common, survey designers may specifically want to ask if a child is adopted.

Now consider the issue of children who are not household members, whether minors or adult children. In many developing countries it is common for parents to send their children to live with other families, particularly for the purpose of sending them to school. While children in this category will be sampled in the households where they currently live, it is often useful to link them to their parents' households for purposes of analysis. This can be done by asking all adult household members whether they have any children who live away from the household. It is also useful to gather information about any adult children who live away from home. This is useful for studying poverty because adult children may support their elderly parents even when they live in a separate household. As a result, it is useful to know how many potential sources of support each household has in the form of adult children living away from home.<sup>3</sup>

Finally, along the same lines, it may be useful to collect some information on the absent spouses and siblings of adult household members, because such spouses and siblings can often be counted on to support the household in times of need. This information is relatively easy to collect for absent spouses, since most households will not have any absent spouses and few households will contain more than one person with an absent spouse. In contrast, the number of potential siblings is very large. One option is to ask only about the number of siblings alive, perhaps distinguishing between men and women (as this may affect the probability of receiving support). Only a few surveys have attempted to collect information on each sibling of all household members; one example is a survey done in Cartagena, Colombia in 1982 (see Bamberger, Kaufmann, and Velez 1984). The basic approach is to ask about such people in a survey form similar to the one on nonresident children in Volume 3 (Part C of the Household Roster module). The two main problems are that the list of people could be very large, especially for households with several adult members, and that people could be siblings of more than one household member. However, neither problem is insurmountable.

### **Draft Module**

This section briefly introduces the draft household roster module. (The module itself is provided in

Volume 3.) The draft module is composed of four parts. Part A is the list of household members. It consists of two pages. The first page provides instructions to the interviewer, along with some questions to be asked of the head of household. The second page contains the actual roster. Part B collects information about the parents of household members—and a small amount of information on the siblings of adult members. Part C collects data on children of household members who do not reside in the household, and Part D collects more detailed information on the siblings of adult household members.

Like many of the other modules in Volume 3, the household roster module has three versions: short, standard, and expanded. The short version is simply Part A by itself. The standard version consists of Parts A, B, and C. The expanded version consists of all four parts, and also provides additional questions about nonresident spouses that should be added to Part A.

### **Notes and Comments on the Draft Modules**

To help the interviewer correctly record all the information in each of the other individual-level modules in the household questionnaire, the part of the roster that contains each individual's name and sex (Questions 1 and 2 of Part A) should be visible and aligned with each individual-specific grid in those modules. (See Chapter 3 for a discussion of ways to design a fold-out roster page.) It is also useful to have two extra columns immediately to the left of Questions A1 and A2 that can be used to write each person's age (in years) and indicate whether the person is a household member. This is useful because the individual-level sections that follow apply only to household members, and in some cases they apply only to persons of certain ages.

**A7.** For some studies of migration it is useful to add the year of marriage. Such a question should come immediately after Question A7, and should be asked only of persons currently married.

**A8.** If the spouse does not live in the household, one may want to ask some questions about his or her whereabouts, since absent spouses can be a source of support in time of need. Such questions should be added for the expanded version of the household roster.

ter. They should be asked immediately after Question A8. The questions to ask are the same as the questions for nonresident siblings in Part D of the household module, particularly: age, years of schooling and degree completed, current employment or other activity, and current location. Another possible question is when the person left the household (to be asked of individuals who were previously members of the current household).

**A10.** If ethnicity varies within a substantial fraction of the households in the country, the ethnicity question in the Household Identification and Control Information page of the metadata module (see Chapter 4) should be removed, and a question about ethnicity should be added immediately after Question A10. The same point applies to religious affiliation.

**A11.** It is best to ask about the number of months spent away from the household because this type of question is usually easier to answer than a question about the number of months during which the member was present.

**A12 (INSTRUCTIONS PAGE).** As explained in the text, other excluded categories (beyond the head of household and newly born infants) could also be treated as household members. Such categories might include recently demobilized soldiers or new spouses of household members.

**PART B.** Survey designers may want to ask a question regarding where each parent currently lives. They may also want to ask the child's age when he or she first lived away from his or her parents. This should be asked separately for each parent.

**B1, B8.** The decision about whether to treat adoptive parents as the real parents will depend on country-specific factors. See the text for further discussion.

**C13.** This question should include codes for foreign countries if it is common for children to live overseas. This was the case in Vietnam.

**PART D.** In addition to being the brother or sister of a household member, this person could also be a parent or child of a household member. For exam-

ple, a child who lives with her uncle or aunt will have already provided some information on his or her parents—one of whom is the sibling of the uncle—in Part B. Or a person 21 or older who lives with his or her parents may have a sibling whose existence has already been recorded in Part C. In such cases one may not want to repeat the information, so an extra question should be added after Question D4 asking if this person is a parent or child of a household member. If the answer is “YES,” the ID code of the person as indicated in Part B or Part C should be provided here, and then the interviewer should go on to the next person. However, in some cases one may want to ask more about the person, in which case all of the questions should be asked of such persons.

**D1.** It is possible for a nonresident sibling to be the brother or sister of more than one household member. In such cases that sibling's name should be written down only one time, and the ID codes of brothers and sisters in the household to whom he or she is related should be recorded in Question D5.

**D10.** The codes here should include countries if it is common for siblings to live overseas.

**D12.** In addition to some basic occupation codes, the following “activity” codes could be added: retired, housewife, student, unemployed, unable to work.

## Notes

The author would like to thank Jere Behrman, Margaret Grosh, and Courtney Harold for comments on a previous draft of this chapter.

1. Not every country follows this standard definition. In some European household surveys the requirement that household members eat at least one meal together each day is dropped; other surveys require some kind of kin relationship among the household members. Thus there is some room for flexibility when defining household membership in specific countries.

2. One possible exception to this is if an analyst is trying to construct a panel data set by matching names (and other information) between different surveys. A well-planned panel survey should not have to resort to this crude and error-prone method. For further information see Chapter 23 on panel data.

3. Whether any support is ever received from these children is a separate question. This type of data is collected in the transfers and other nonlabor income module (which is introduced in Chapter 11).

## References

- Bamberger, Michael, Daniel Kaufmann, and Eduardo Velez. 1984. "Research Methodology: Design and Analysis Issues." World Bank, Water Supply and Urban Development Research Department, Washington, D.C.
- United Nations. 1989. "Household Income and Expenditure Surveys: A Technical Study." National Household Survey Capability Programme, New York.



# 7 Education

*Paul Glewwe*

In recent years a consensus has developed among agencies and individuals working in economic development that investments in human capital, particularly investments in education, are crucial for economic growth (World Bank 1990; UNDP 1990; Becker 1995). Yet many developing countries continue to have serious problems with their educational systems, and many observers argue that the provision of education in most developing countries is highly inefficient (Lockheed and Verspoor 1991; Hanushek 1995).

In order to devise policies that improve their educational systems, policymakers in developing countries need accurate information on education. However, in many countries the only data that policymakers have at their disposal is a small amount of information collected from public schools. These data give only a partial picture of how, and how successfully, students are educated. School-based data provide no information on children who do not attend school or on what happens to students after they leave school.

Household surveys can fill this gap by supplying basic descriptive information on which children go to school, the characteristics of the schools they attend, how long they attend, the costs of their schooling, and what happens after they leave school. Such information provides a better foundation for research on how to improve education systems in developing countries. Household surveys can also provide policymakers with important information on the impacts of specific government policies.

However, household surveys can perform these important functions only if they are well designed. The purpose of this chapter is to explain how to design

LSMS and other multitopic household surveys to collect data that can be used to investigate a wide variety of education policy issues in developing countries.

The first section of this chapter reviews major policy issues on education. The second section examines how data on education from household surveys can be used to address these issues. The third section introduces draft modules for collecting data on education in multitopic household surveys. (The modules themselves are presented in Volume 3.) The fourth and final section provides explanatory notes on the draft modules introduced in the third section.

## **Education Policy Issues in Developing Countries**

Policymakers who work on education issues need information on the current situation and estimates of how the current situation would change in response to changes in government policy. More specifically, they need accurate information on educational outcomes, such as school attendance and skills learned, and they need to know what impact government education policies have on each educational outcome. In addi-

tion, policymakers would like to know the impact of these educational outcomes on other socioeconomic outcomes, including income, migration, and health status.

LSMS and other multitopic household surveys can provide data that meet all of these needs. In such surveys the education module of the household questionnaire is the main source of information on educational outcomes. This module is also an important source of information on many of the determinants of educational outcomes. Information on other socioeconomic outcomes is collected in other modules of LSMS-type surveys.

This section reviews the most important policy issues concerning education in developing countries, focusing primarily on the impact of government actions on educational outcomes. A final subsection discusses whether the education module should collect additional data to investigate the impact of education on other socioeconomic outcomes.

### *The Basic Educational Outcomes*

There are several basic educational outcomes of interest to policymakers. In general, a child or young adult enrolled in school acquires:

- Basic cognitive skills such as literacy and numeracy.
- Complex cognitive skills such as reasoning ability.
- General knowledge on a wide variety of subjects, which may include science, geography, agriculture, and health.
- Specialized technical skills (beginning at the secondary level).
- Diplomas and certificates attesting to the completion of specific levels of schooling.
- Values and behavioral norms, both those that are part of the curriculum (for example, good citizenship) and those that arise through social interaction among students (for example, aspirations to a “high status” occupation).

Measurement of learning outcomes (in other words, measurement of the acquisition of cognitive skills, general knowledge, and specialized technical skills) can be complicated. Measuring values and norms can be even more difficult. One way to avoid difficulties in measuring learning outcomes is to collect data on other variables that are closely related to learning. For example, it is much easier to collect data on years of schooling completed and on certificates or diplomas received, and these data should be highly correlated

with skills and knowledge. It is also likely that years of schooling and diplomas received are correlated with the values and norms acquired in school, especially the values that are explicitly part of the curriculum.

The six types of educational outcomes listed above can be thought of as the main “outputs” of the education process. Other schooling variables can be thought of as “inputs,” including years of schooling, grade repetition, daily attendance, and household expenditures on school supplies. The number of years of schooling is perhaps the most important input, but daily school attendance clearly matters as well. Repetition indicates whether sufficient learning is occurring and also affects learning directly; thus it can be thought of as both an output and an input. Finally, household expenditures on schooling consist of school fees and the amounts (and prices) of books and other school supplies purchased by the household.

In this chapter, “educational outcomes” will be used to refer to both outputs and inputs. Thus the following schooling variables comprise the basic set of educational outcomes of interest to policymakers: basic cognitive skills, complex cognitive skills, general knowledge, specialized technical skills, diplomas and certificates, values and norms, years of schooling completed, daily attendance, grade repetition, household expenditures on schooling, and current enrollment. Current enrollment is technically neither an input nor an output, but it is important because it allows analysts to distinguish children whose formal education is finished from children who are still in school.

Government policies affect educational outcomes by influencing their determinants. The determinants of educational outcomes can be classified into three types: child characteristics, household characteristics, and school characteristics. In virtually every developing country, most schools are run by the government, which means that government policies have a major impact on the characteristics of schools. (Child and household characteristics may also be affected by government policies, as discussed below.) Government policies can be divided into those that affect prices and those that affect other school characteristics. This distinction is important because prices affect learning only indirectly by influencing inputs chosen by the household, such as the number of years a child is enrolled in school. In contrast, most other school characteristics (for example, basic classroom materials such as blackboards and desks) have a direct effect on learn-

ing. These other characteristics of schools can also have indirect effects because families may change other inputs, such as years of schooling, in response to changes in school characteristics.

#### *Government Policies Regarding the Price of Schooling*

The total cost to a household of enrolling a child in school is the sum of the direct money costs and the opportunity costs. The direct costs include tuition and other required fees—parents' association fees, examination fees, sports fees, special fundraising levies—as well as expenditures on uniforms, textbooks, other learning materials (pencils, exercise books), transportation, meals at school, and in some cases lodging. If a student receives a scholarship or voucher, its value can be subtracted from the household's direct costs. Opportunity costs are the implicit costs of the time that children devote to schooling, including the time they spend in the classroom, traveling to school, and doing schoolwork at home. If a child moves away from home to attend school, the household typically loses the use of all of his or her time. In developing countries children's time is valuable because they often help with household chores, work on household agricultural land, assist in operating nonagricultural household businesses, and even work for wages.

The discussion so far has been in terms of "costs," not "prices." In general, cost equals price times quantity. Several different prices apply to schooling. The first is the mandatory tuition fee per year (or per term) of schooling. Other prices include the prices for the various learning materials that parents are expected or even required to purchase, such as uniforms, pencils, exercise books, and, in some countries, textbooks. Then there are prices for transportation, meals, and lodging. In addition, there is the price of an hour or day of a child's time, which determines the opportunity costs incurred by households that send their children to school. While tuition fees must be paid if a child is to attend school, other schooling costs may be optional.<sup>1</sup> For example, in many developing countries where parents are expected to purchase textbooks for their children, many parents do not do so or purchase only some of the required books, yet still enroll their children in school. Also, although official policy often requires children to wear school uniforms, this is not always enforced in practice.

In general, households cannot usually change the prices that they face. However, by making decisions

regarding optional expenditures on schooling, they can at least partially control the total costs of schooling (henceforth referred to as school expenditures). As will be seen in the second section of this chapter, the distinction between optional and required expenditures is important when discussing what data are needed for analyzing policy issues.

Many education policy issues in developing countries concern school prices. The levels at which tuition and other required fees are set is a matter of constant debate. Some countries have provided free public education at virtually all levels for many years, while other countries set public school fees at a level high enough to reimburse the government for much of the cost it bears from providing education services. Policymakers face two opposing pressures in setting school fees. On the one hand, increases in school fees tend to reduce enrollment and eventual school attainment, especially among poor households. On the other hand, budgetary resources in most developing countries are scarce, which puts pressure on governments to raise fees to fund the operation of public schools. This creates a fundamental dilemma for policymakers; there are no easy answers, particularly when reducing poverty is a key concern.

Many proposals for improving education in developing countries focus on pricing policies. Some education experts have argued that public school fees should be very low at the primary level and perhaps the lower secondary level but should be relatively high for upper secondary and postsecondary schooling (World Bank 1995). This would increase equity by raising primary school enrollment, particularly among poor households, and by reducing subsidies to better-off households, whose children are most likely to enroll in higher education. If the social returns to primary school are higher than the private returns and there is little difference between the social and private returns to higher education, this would also be more efficient.

Other experts recommend increasing public school fees but using the funds to improve the quality of schooling provided, since there is evidence that many households, even poor ones, are willing to pay more for higher-quality education services. A third possibility, often referred to as targeting or price discrimination, is to raise fees at all levels of education but to reduce the fees for poor households in order to encourage them to send their children to school. A

fourth suggestion is to set fees at fairly high levels while providing loans to some or all students to ensure that credit constraints, which are presumably more common among the poor, do not prevent parents from sending their children to school. Finally, some have argued that government-run schools are inherently inefficient and that private schools should be promoted, which could be done by providing families with vouchers that they could use to send their children to either type of school (West 1996).

The validity of the arguments in favor of these different policies on school fees depends on how households are likely to react to each policy option. Thus it is important that the education module in multitopic household surveys be designed to gather data that allow analysts to estimate behavioral models of the educational choices that households make.

There are other prices that parents face when sending their children to school, and government policies can affect these prices. The price at which uniforms are available can affect children's educational outcomes, particularly if government schools require uniforms. Similarly, if the school does not provide textbooks, textbook prices will also affect households' schooling decisions. If textbook prices are high, parents may withdraw their children from school or send them to school without a full set of textbooks. The same applies to other learning materials that parents are expected to purchase. Distance can also be viewed as a price; schools that are far from a household's dwelling may discourage parents from enrolling their children in school because of the high opportunity costs of the children's time spent traveling to school—and in some cases due to direct transportation costs. In many developing countries, schools are located quite far from many rural communities, especially at the secondary level. For example, in rural areas of Vietnam, the distance to the nearest upper secondary school was more than 10 kilometers in 25 percent of the rural communes sampled. Proposals for providing student loans can also be treated as price policies, since they aim to provide more children with access to schooling rather than to change school characteristics.

This discussion of pricing policies suggests that household surveys should be designed to answer the following four questions:

- How would educational outcomes be affected if mandatory school fees were changed?
- How would educational outcomes be affected if the prices of items that parents are expected to pur-

chase, such as uniforms, textbooks, pencils, and exercise books, were changed?

- How would educational outcomes be affected by changes in the distance from households to the nearest school?
- How would educational outcomes be affected by a student loan program?

Several points need to be made regarding these four questions. First, the question about changing mandatory school fees includes policies on vouchers and scholarships, since providing a voucher or scholarship changes the effective mandatory school fee. Second, to address the issue of how increases in school fees accompanied by school quality improvements affect educational outcomes, it is necessary to combine the answer to the first question with estimates of the impact of school quality on educational outcomes.

Third, when the purchase of uniforms, textbooks, and other learning materials is required to enroll a child in school, researchers may be tempted to treat the prices of these commodities as additional mandatory school fees. However, this is often not advisable. For example, uniforms usually last more than one year, so their price is not an annual price. Also, both textbooks and uniforms can be passed from older children to younger siblings, so again the price is not necessarily annual. In general it is best to treat the prices of learning materials as separate variables rather than adding them to tuition to obtain a single price variable.

Fourth, measuring the distance from a household to the nearest school can be complicated by various factors. There may be more than one school from which the household can choose. Some families may decide to move to a new dwelling in order to be nearer to a certain school. Other families may send their children to live away from home in order to attend a particular school. Finally, the availability and cost of public and private transportation can alter the impact of distance on educational outcomes. In each case distance is not a pure price but may be partially under the household's control. The survey design problems presented by these issues of distance will be discussed further in the second section of this chapter.

#### *Government Policies Regarding the Characteristics of Schools and Teachers*

There are many important policy issues concerning school and teacher characteristics in developing countries. (For convenience, the term "school characteris-

tics” will be used in the rest of this chapter to refer to nonprice school characteristics.) It is useful to divide these characteristics into two types. The first type is characteristics that concern what actually happens in the classroom. These can be thought of as school quality variables. School quality can be defined as all of the characteristics of classrooms, the teachers in them, and the teaching methods used that directly affect how much children learn and what values they acquire. Tuition fees and other prices are not considered components of school quality because, although prices may be correlated with quality, they do not directly affect learning.<sup>2</sup> Even though this definition of school quality focuses on learning (the acquisition of skills and knowledge) and values, it affects all of the educational outcomes discussed at the beginning of this section. For example, parents are likely to keep their children in school longer if school quality improves (assuming that other factors, such as school fees, do not change) because the benefit of a year of schooling rises while its price remains the same. The many different aspects of school quality can be divided into two categories:

- Material inputs in the classroom, such as blackboards, textbooks, and the physical condition of classrooms.
- Teacher characteristics and the pedagogical practices teachers use.

The second type of school characteristics is those concerned with school management and school policies, which affect student learning indirectly by determining what takes place inside the classroom. These also can be divided into two categories:

- School management variables, which refer to the managerial structure of the school, including the characteristics of the school principal and the overall system of incentives for teachers and other staff.
- Admissions and advancement policies, which determine the schools and classes students may attend.

Material inputs range from the most basic school supplies, such as chalk, blackboards, textbooks, and desks, to much more expensive and sophisticated pedagogical tools, such as personal computers. The poorer developing countries lack even the most basic material inputs (see Glewwe, Kremer, and Moulin 1999 on Kenya and World Bank 1997 on India). By far the most important policy question in these countries in this context is: which material inputs are most cost-effective?<sup>3</sup> In other words, which inputs bring about

**Box 7.1 Using a School-Based Sample to Examine the Impact of School and Teacher Characteristics**

Much research on the impact of school and teacher characteristics on educational outcomes, particularly on test scores, is based on data collected from a sample of schools rather than a sample of households. Such data typically contain detailed information on schools but only a small amount of information on schoolchildren’s households. (This is because the information on each child’s household is obtained by asking the child, not by visiting the household.) The lack of data on the children’s households can lead to serious estimation problems. If some children do not attend school, as is often the case in developing countries, it is best to use a sample of households rather than a sample of schools to collect education data. Ideally, household surveys should collect data from both schools and households. Such data are essential for the analysis of certain policy questions, such as the impact of school and teacher characteristics on the likelihood that children will be enrolled in school.

the largest improvement in educational outcomes per dollar spent? Answering this question requires information on how each material input affects educational outcomes and on the cost of each input. This information is useful not only for basic inputs but also for more sophisticated ones.

There are a large number of policy issues regarding teachers and how they teach. Among the most important are:

- The effect of student-teacher ratios on educational outcomes.
- The impact of teacher training and the general educational level of the teacher (including teacher knowledge as measured by test scores) on educational outcomes.
- The role of teacher motivation and morale.
- The impact of various pedagogical techniques, such as distance education (radio instruction), on learning.<sup>4</sup>
- The role of female teachers in encouraging parents to educate their female children.

Some of these issues involve factors that are relatively easy to measure (such as student-teacher ratios, the extent to which a teacher has been trained, and the sex of a teacher), while others are more difficult to observe (such as teachers’ motivation and the pedagogical practices teachers use). A similar observation can be made regarding the costs of teacher characteristics and

pedagogical practices. Some costs, such as raising teachers' pay to increase their morale, are relatively easy to measure, while others, such as switching to a new pedagogical technique, are harder to measure. However, in principle, any change involves a cost, and the changes that bring about the greatest improvements in educational outcomes per dollar spent should be given the highest priority.

Key policy issues concerning systems of school management include:

- The impact of the education levels, training, experience, and management styles of school principals on educational outcomes.
- The effect of decentralized decisionmaking (allowing individual schools and teachers to choose curricula and allocate resources) on educational outcomes.
- The relative efficiency of public and private schools.
- Methods by which principals can motivate teachers and increase parental participation in schools.

Finally, the current issues regarding school admissions and advancement policies are: the minimum age of enrollment into first grade; whether area of residence should determine which public school a child can attend; policies on grade repetition; and standards for determining whether a student is allowed to advance to the next grade or level of schooling.

To assess the cost-effectiveness of both school quality initiatives and practices regarding management, admissions, and advancement, policymakers need information on the impact of changes in school characteristics on educational outcomes as well as information on the costs of those changes. In general, household surveys can only provide answers on how changes in school characteristics affect educational outcomes; most information on the costs of changes in school characteristics cannot be collected using a school questionnaire in a multitopic household survey and thus must be collected in a separate data collection exercise. In particular, a comprehensive study of actual resource costs is needed, which can be quite complex. For an example see Ilon (1992).

In summary, household surveys should be designed to provide answers to the following four questions:

- What impact do different material inputs have on learning and other educational outcomes?
- What impact do teacher characteristics and specific teaching practices have on learning and other educational outcomes?

- What impact do different types of school management policies have on learning and other educational outcomes?
- What impact do different admissions and advancement policies have on learning and other educational outcomes?

In general, households cannot provide reliable information on school characteristics, since it is unlikely that they will have detailed knowledge of them. Instead, this information should be collected in the community questionnaire, or by adding a school questionnaire to the survey. Many past LSMS surveys collected community data, but only three (Ghana, Jamaica, and Vietnam) attempted to collect data from local schools. (The Morocco LSMS survey gathered detailed data on schools as part of the general community questionnaire.) Practical guidelines on how to collect school data will be provided later in this chapter. The point to bear in mind at this stage is that collecting data on school characteristics is not simply a matter of modifying the household questionnaire.

#### *Other Government Policies That Affect Educational Outcomes*

Several government policies not usually thought of as education policies can have important effects on educational outcomes. Perhaps the most important of these are policies that affect child nutrition. There is ample evidence that children's nutritional status, especially in the first one or two years of their lives, can have a sizeable impact on later school performance (Glewwe and Jacoby 1995; Glewwe, Jacoby, and King Forthcoming; Alderman and others 1997). Therefore, all government policies aimed at improving children's health and nutrition, such as immunization campaigns, nutrition education programs for young mothers, sanitation programs, school feeding programs, early childhood nutrition initiatives, and food stamp programs, have potential implications for educational outcomes.

A second set of policies that can affect educational outcomes are those that influence the opportunity cost of children's time, such as policies regarding child labor, rural infrastructure, agricultural extension, and childcare facilities. For example, providing a local source of potable water to a rural community may reduce the opportunity cost of children's time, particularly for girls, who may no longer need to walk long distances to obtain water for their families. Of course, some policies can increase the opportunity costs of

schooling. For example, providing access to a new agricultural technology may raise the productivity of child labor on family farms.

Another way in which government policies can affect educational outcomes is by changing the returns to schooling in the labor market. For example, a government that decides to reduce its budget deficit by imposing a freeze on hiring new secondary school or university graduates will decrease the private benefits of schooling (particularly if government jobs pay higher wages to well-educated workers than do comparable private sector jobs). Such a policy will tend to reduce school attainment.<sup>5</sup> Alternatively, allowing for direct foreign investment may increase job opportunities for educated workers, leading to an increased demand for education.

This means that household surveys should gather data that can help answer the following questions:

- What is the impact of child nutritional status on educational outcomes?
- What impact do child wage rates and child productivity in self-employment and household chores have on educational outcomes?
- How do labor market conditions affect educational outcomes?

#### *Policies on Vocational Training and Postsecondary Education*

The discussion so far has implicitly focused on primary education and general (as opposed to specialized) secondary education. Yet almost all developing countries also provide technical training and postsecondary education. In many developing countries it is difficult to use household surveys to collect data on these types of educational services because very few people receive technical training or study at the postsecondary level. Thus those who do receive vocational training or postsecondary education are unlikely to appear in the sample in sufficient numbers to make it possible to draw reliable conclusions from the data. Nevertheless, a brief review of these policies is worthwhile.

Many policy issues at the postsecondary level involve pricing, and the discussion above on pricing policies still applies. Several of the policy issues with respect to vocational training are discussed in this book in Chapter 9 on employment (and therefore need not be discussed here). However, there are two general questions regarding both postsecondary education and vocational training that have specific impli-

cations for the design of the education module of LSMS and other multitopic household surveys:

- Who receives this type of education?
- What are the benefits of this type of education?

The first question implies that the education module should be designed to collect information on vocational training and other “nongeneral” education, such as apprenticeships,<sup>6</sup> and on postsecondary education, including the subject studied. The second question pertains to the impact of education on socioeconomic outcomes, which will be discussed in the next section.

#### *Policies Concerning the Effect of Educational Outcomes on Other Socioeconomic Outcomes*

The discussion so far has focused on the determinants of educational outcomes, but an equally important use of data from the education module of a multitopic household survey is analysis of the impact that educational outcomes have on other socioeconomic factors. This raises the question: what kind of educational outcome data should be collected in the education module to support these kinds of analyses? The indicator most commonly used in studies of the impact of education on other socioeconomic outcomes is years of schooling, which has been collected in all past LSMS surveys. In addition, almost all past LSMS surveys have collected information on certificates or diplomas obtained by household members, which household members are currently enrolled in school and, for those who are enrolled, household expenditures on education. However, most previous LSMS surveys have not collected data on basic cognitive skills, general knowledge, specialized technical skills, values, or grade repetition. Also, only about half of these surveys have collected any information on daily attendance (and these have usually asked only about attendance during the seven days prior to the interview).

There is one educational outcome that has usually not been collected in past LSMS surveys but could be collected relatively easily in future surveys: grade repetition. This information is useful for computing actual years of school attendance, which in turn can have implications for estimating rates of returns to schooling (Behrman and Deolalikar 1991). Therefore, this variable should be routinely collected in surveys in those countries where grade repetition is common. Information on daily attendance over a short period of time (such as the past one or two weeks) would also

be reasonably easy to collect. However, this information is less important for understanding the impact of schooling on socioeconomic outcomes because attendance is primarily an input into other outcomes (such as skills learned or diplomas received), instead of being an output in its own right.

Thus the major question is whether information should be collected on skills, knowledge, and values. This will be discussed fully in the second section of this chapter. However, four comments can be made at this stage. First, as discussed above, collecting information on skills is a major undertaking. Collecting information on values is probably easier, but is rarely done in nationally representative surveys in developing countries. Second, information on the skills, knowledge, and values of adults can be extremely useful for examining the impact of education on various socioeconomic outcomes. (See Murnane and others 1995 for an example from a developed country and Knight and Sabot 1990 for an example from a developing country.)

Third, information on children's skills can be used to investigate the determinants of other socioeconomic outcomes using an indirect approach.<sup>7</sup> For example, it would be possible to examine how education policies affect adult socioeconomic outcomes by combining estimates of the impact of education policies on children's skills with estimates of the impact of the skills of adults on the socioeconomic outcomes of adults. (See Glewwe 1999 for an example.) Fourth, in school systems in which skills are highly correlated with years of schooling, there is little reason to collect skills data. School systems in which children must pass standardized examinations to continue to the next grade are particularly likely to produce such a high correlation. However, the school systems in many, if not most, developing countries do not operate in this manner, which implies that there is a strong argument for collecting data on skills, both for adults and for children, in such countries.

One line of research of recent years that deserves specific discussion is the impact of school quality on wages and other labor market outcomes. After controlling for years of schooling, do individuals who attend relatively high-quality schools have higher wages and other desirable labor market outcomes? The discussion of school quality data in the previous subsections focused on current school quality, but researchers need data on past school quality to answer

this question. In general there are only two ways to collect information on past school quality: asking respondents to recall characteristics of their schools and finding data on past school quality from some other source. In either case the chance of obtaining useful data are small. Respondents' recollections on the characteristics of the schools that they attended decades ago are likely to be highly unreliable, especially given the complex nature of school quality. Similarly, in the vast majority of developing countries there is unlikely to be much useful information on school quality in previous decades, and the few who do have such data probably only have a small amount of information for each school. However, one approach to this issue, pursued by Glewwe (1999), is to examine the impact of current school on cognitive skills, and then to examine the impact of those skills on labor market outcomes. This was done using the data from the 1988–89 LSMS survey in Ghana. The only alternative is to begin collecting panel data today and follow individuals for decades, which implies a wait of many years before results can be obtained. Moreover, it may be difficult to collect panel data in developing countries; see Chapter 23 for a discussion of collecting panel data as a part of LSMS surveys.

Finally, analysts may want surveys to collect data on variables that predict educational outcomes, such as years of schooling. The reason for this is that estimation problems can arise if an educational outcome variable is used instead of its predicted value.<sup>8</sup> Some "predicting" variables are likely to be collected already in an education module that focuses on the impact of various government policies on educational outcomes, as will be seen below. In addition, one generally useful variable for predicting the educational outcomes of an individual, especially of an adult, is the education level of the individual's parents. Data on parents' education should be collected regardless of whether the parents are members of the individual's household or even whether they are alive. Such data have often been collected in past LSMS surveys, usually in the household roster. See Chapter 6 for a discussion of the household roster module.

### **Data Needed to Address the Major Policy Issues**

This section discusses the data required to address the policy issues raised in the previous section. It also dis-

cusses some methodological issues concerning the estimation of relationships that provide answers to policymakers' questions, although this discussion has been kept to a minimum. Before turning to the issue of what data are needed, it is useful to review briefly, in nontechnical terms, what determines educational outcomes. This will provide a framework for comparing the advantages and disadvantages of collecting different kinds of data.

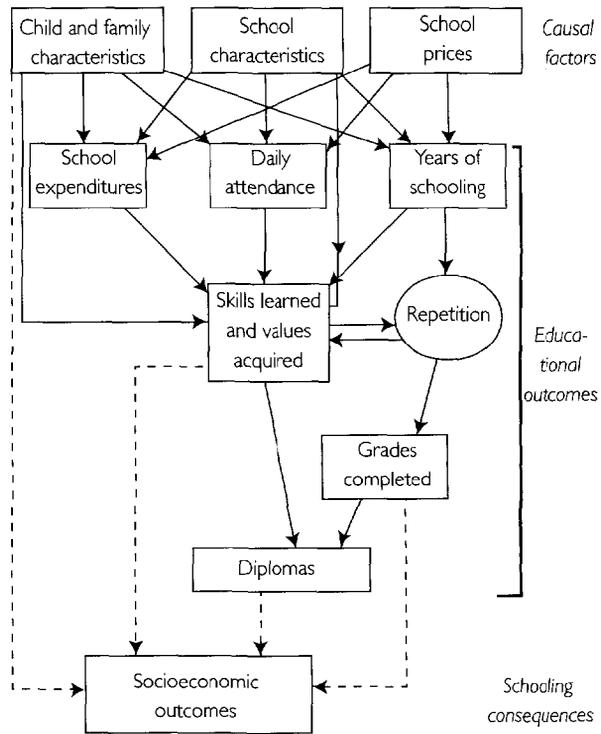
**The Determinants of Educational Outcomes**

Figure 7.1 provides a visual framework of the determinants and consequences of educational outcomes. Although these could be presented as mathematical equations, most survey designers probably prefer a visual representation. This subsection describes the "model" on which this figure is based.

At the top of Figure 7.1 are three boxes, labeled child and family characteristics, school characteristics, and school prices. These three kinds of causal factors ultimately determine all educational outcomes. School characteristics and school prices were described in detail in the previous section. The characteristics of a child's family that determine his or her educational outcomes are: the education of the child's parents; the household's wealth, income, or both; the value that the child's parents place on schooling (including attitudes towards educating girls); the existence of family farms or businesses (which may affect the opportunity costs of a child's labor); the size of the family; and the family's access to credit. The characteristics of a child that affect his or her educational outcomes are the child's "learning ability" (which includes both genetic and environmental factors), sex, motivation for schooling, position in the birth order, and past and present nutritional status. These three groups of causal factors influence three fundamental educational outcomes that are directly controlled by parents (and sometimes their children): household expenditures on schooling, daily attendance, and years of schooling. These causal relationships are indicated by the arrows in Figure 7.1.

One level down in Figure 7.1 are skills learned and values acquired. All of the three educational outcomes that parents can directly control—school expenditures, daily attendance, and years of schooling—are important determinants of skills and values. Child and family characteristics and school quality also directly determine skills and values. These causal influences are indicated by arrows in Figure 7.1.

**Figure 7.1 Determinants and Consequences of Schooling Outcomes**



Source: Author's summary.

Two points are particularly notable. First, there are no arrows from school prices (both mandatory school fees and the prices for various learning materials) to skills learned or values acquired because the former have no direct effect on the latter. They only have an indirect effect through years of schooling, school expenditures, and (possibly) daily attendance. Second, the five groups of factors that directly affect skills learned—child and family characteristics, school quality, school expenditures, daily attendance, and years of schooling—can be thought of as a "production function" for skills learned in school (Hanushek 1986). This concept of a production function, discussed in the next subsection, has had a strong influence on studies of education done by economists. In principle, a similar "production function" exists for values acquired, but little research has been done on how schooling affects values in developing countries.

The next level of Figure 7.1 has only one educational outcome: grades completed. If children never repeated grades, there would be no need for this level, because grades completed would equal years of schooling. However, grade repetition is common in

many developing countries (Lockheed and Verspoor 1991). Perhaps the most important determinant of grade repetition is skills learned, because many schools do not allow a child to advance to the next grade until a certain degree of proficiency is reached in various skills. Thus the figure for the number of years repeated, which is the difference between years of schooling and grades completed, depends on skills learned. Repetition can also affect skills learned since some classroom time may be redundant when a child repeats a grade. This is indicated in Figure 7.1 by the small arrow going from repetition to skills learned and values acquired.

The last level of educational outcomes in Figure 7.1 shows that schools grant diplomas and certificates to students who successfully complete a certain grade or pass a certain examination (hence the arrow from skills learned to diplomas). In some countries, people who complete a grade but do not pass the associated examination are officially considered as having completed that grade. In these countries it is important for survey questionnaires to distinguish between finishing the grade and passing the examination.

At the bottom of Figure 7.1 are socioeconomic outcomes (which include employment, wages, farm productivity, health status, fertility, migration, and the nutritional status of children). These data come from other modules of a multitopic household survey questionnaire. The dashed lines in Figure 7.1 show the causal relationships that determine these socioeconomic outcomes.

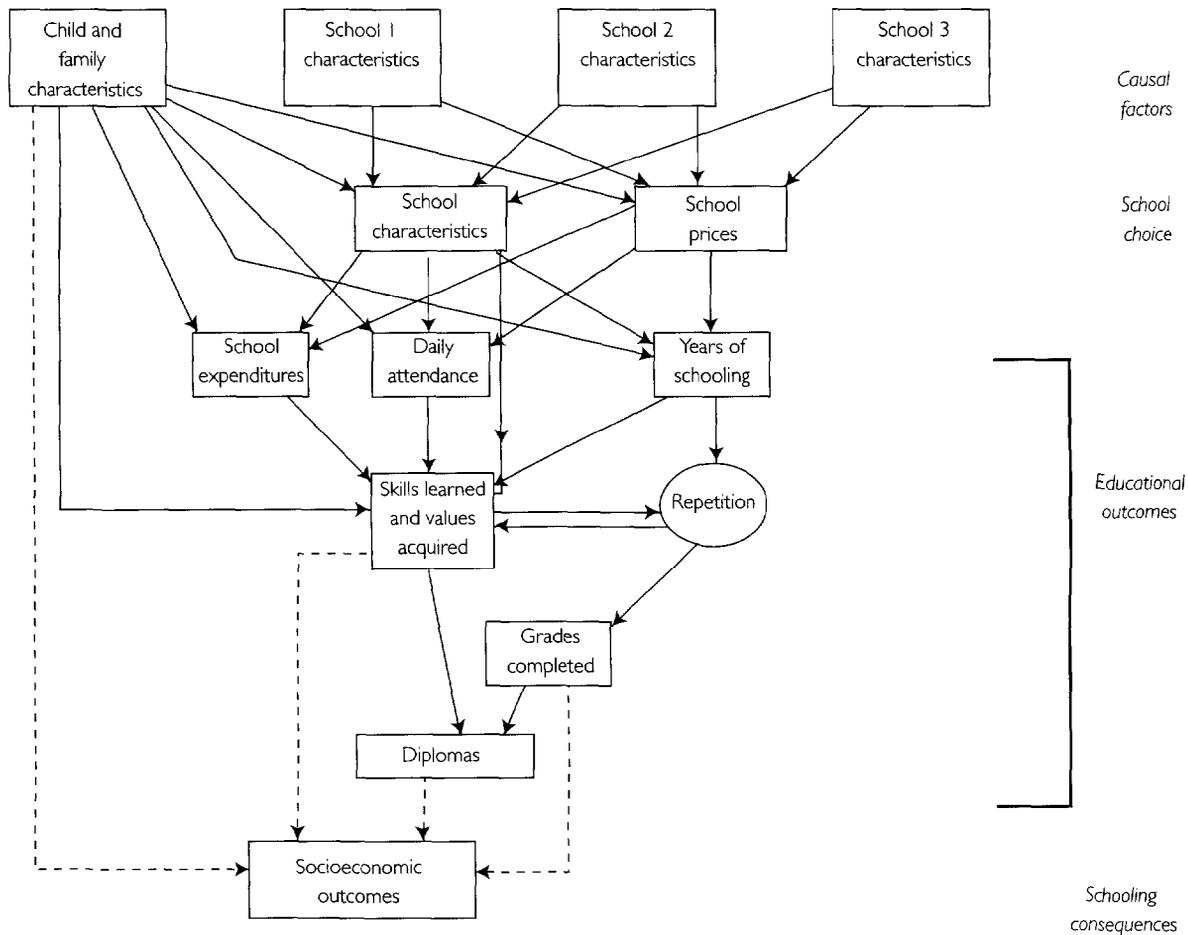
Figure 7.1 assumes that school characteristics and school prices are exogenous—that is, beyond the control of the household. However, in many developing countries, parents have several schools from which to choose. Even when only one local school exists, parents may decide to send their child to a boarding school or to live with relatives located near a “good” school.<sup>9</sup> This implies that school characteristics and school prices are not exogenous but are determined by the parents when they select a school for their children. However, the schooling options themselves—in other words, what local schools, and possibly boarding schools, are available—are generally beyond the control of the parents. (In some cases, even the choice of schools may be endogenous. This is discussed in later subsections.) Thus the characteristics of each schooling option do not depend on a choice made by the household. Figure 7.2 demonstrates this situation. Figure 7.2 is the same as Figure 7.1 except that the boxes for

school characteristics and school prices are no longer at the highest level of the diagram. Child and family characteristics remain at the highest level and are still causal factors, but the other causal factors are the characteristics of three schools from which the family can choose. (The number 3 is purely illustrative; actual numbers of choices vary by country and by household.) The boxes on the characteristics and prices of the school attended are labeled to show that these variables are determined by households’ choices.

### *Some Basic Econometric Concepts*

The rest of this section discusses what data are needed to provide answers to the policy questions in the first section of the chapter. Some reference to research methodologies is inevitable in this discussion. Therefore, this subsection provides a brief, nontechnical discussion of three different econometric relationships that researchers may want to estimate. For a fuller and more technical discussion of basic econometric concepts see Chapter 26.

Household survey data can be used to estimate three types of socioeconomic relationships. The first is a reduced form relationship, in which an educational outcome is determined by exogenous, causal factors.<sup>10</sup> These causal factors are a subset of the child, family, and school variables (characteristics) discussed in the previous subsection. Some child and family variables are generally accepted by economists and other social scientists as exogenous. Other such variables are generally accepted as being endogenous. Finally, there are some variables for which no consensus has been reached. School variables fall into the third category; whether they are exogenous or endogenous is a matter of much debate. At the simplest level, there are the two scenarios depicted in Figures 7.1 and 7.2. If there is only one school from which to choose (Figure 7.1), it may be reasonable to assume that the school’s characteristics are exogenous to the household. In Figure 7.2 there are several schools from which to choose. In this case, while the characteristics of the school chosen are not exogenous, the characteristics of the three choices are. However, further problems of endogeneity may arise under each scenario, because households may migrate to live near desirable schools, households may take actions to change local school characteristics, and governments may make decisions regarding public schools based on specific local conditions, which in turn often affect households’ choices. To fully address

**Figure 7.2 Determinants and Consequences of Schooling Outcomes When There Is a Choice of Schools to Attend**

Source: Author's summary.

these potential problems, special data and estimation techniques are required. These will be discussed below.

A second relationship that can be estimated is a production function for educational outcomes. In general, skills learned is the only educational outcome that has been depicted as a production process.<sup>11</sup> The two differences between reduced form estimates and production function estimates (also known as structural estimates) of skills learned are: reduced form relationships use only exogenous variables to explain learning while production functions quite often use endogenous variables, and reduced form relationships include both direct and indirect causal relationships while production functions are concerned only with direct causality. Figures 7.1 and 7.2 illustrate the second difference. The direct causes of skills learned are shown by the five arrows that lead directly to skills learned, while, in contrast, indirect causes affect learn-

ing by first affecting some other variable, such as school expenditures or years of schooling. Because school fees and other prices affect learning only indirectly, there is no direct effect (that is, no arrow going directly from school prices to skills learned), so the learning production function excludes school prices.

It is much harder to estimate production functions than to estimate reduced form relationships. However, the production function is useful for understanding the mechanisms by which school quality improves educational outcomes. Production function relationships are generally assumed to be fairly stable over long periods of time, while reduced form estimates can change as socioeconomic conditions change. This is because reduced form estimates reflect household behavior, while production function relationships are, by assumption, "technological relationships" not altered by household behavior. For example, the

reduced form impact of parental education on learning may diminish as a society becomes wealthier because parental education usually leads to higher incomes, which, in poor societies, implies that parents buy more basic learning materials such as textbooks for their children. As national wealth increases, all parents are able to purchase textbooks, which diminishes this indirect impact of parental education on learning.

The third type of relationship is a conditional demand relationship. It examines the impact of selected exogenous and endogenous variables on educational outcomes. In particular, it concerns the estimation of the determinants of educational outcomes conditional on certain variables of interest that may not be exogenous. Conditional demand relationships can be thought of as an intermediate category between reduced form relationships and production functions. They are not reduced form relationships because they include endogenous variables as causal factors, yet they are not production functions because they include some variables that affect educational outcomes only indirectly.

An example may make this clearer. Suppose a researcher is interested in studying the impact of household income on a particular educational outcome. If household income were exogenous, he or she could simply estimate the reduced form. Yet household income is, to some extent, a household choice, because removing a child from school and putting him or her to work increases household income. Thus one cannot estimate the impact of income as a reduced form relationship. On the other hand, one cannot estimate the impact of income as part of a learning production function, because income does not directly affect learning; its effect is indirect through the goods and services purchased with that income. Conditional demand estimates attempt to measure the impact of an exogenous increase in income. Such an estimate of the impact of income on educational outcomes could be used to assess the impact of economic growth on particular educational outcomes, such as learning or grade attainment. As explained above, this is neither a reduced form relationship nor a production function, but rather a conditional demand relationship.

Conditional demand relationships are, in general, harder to estimate than reduced form relationships but not as hard to estimate as production functions. The difficulty in estimating conditional demand relationships is the need to find variables that predict the

endogenous variables but do not directly affect the educational outcome of interest. (Such variables are called instrumental variables.) The advantage relative to estimating production functions is that the number of endogenous variables in the regression is generally smaller, so fewer instrumental variables are required.

#### *Using Cross-Sectional Data to Investigate the Determinants of Educational Outcomes*

The data collected in most previous LSMS surveys have been cross-sectional data—data gathered in surveys in which households were visited only once. (A household's interview may comprise two or more visits over a week or two, but the resulting "picture" of the household is for a single point in time.) Thus most analyses done on education issues using data from LSMS surveys have used cross-sectional data. This subsection examines how cross-sectional data can be used to understand the impact of government policies on educational outcomes and, more generally, to understand the determinants of educational outcomes. The subsection makes specific recommendations on the data needed to analyze the likely impacts of the different types of education policies discussed in the first section of this chapter.

Before beginning a detailed discussion of data needs, a general point must be made. Accurate estimates of the determinants of educational outcomes require data on all variables that are believed to be causal factors. This is true for all three types of relationships discussed in the previous subsection (reduced form, production function, and conditional demand). Figures 7.1 and 7.2 show that child characteristics, family characteristics, and school variables jointly determine the three educational outcomes most directly under parents' control—school expenditures, daily attendance, and years of schooling. These three educational outcomes, along with child, parent, and school characteristics (but not school prices), then determine all other educational outcomes. Many of these variables will be correlated with each other, and statistical theory shows that estimates of the impact of the different variables based on regression analysis are likely to be biased if causal variables left out of the regression are correlated with causal variables that are included (an estimation problem known as omitted variable bias). Thus, in order to understand fully the determinants of any educational outcome, the researcher must have information on all child, household, and school variables that affect that educational outcome.

A simple example demonstrates this point. Suppose a researcher is trying to estimate the impact of children's nutritional status on educational outcomes, but he or she lacks data on the distance to the nearest school. Estimates done without distance information can yield biased estimates of this impact because children's nutritional status may be correlated with the distance they have to travel to the schools they attend. Households that live far away from schools may also live far away from health clinics. Living far from a school could lead to unfavorable educational outcomes, while living far from a health clinic could lower nutritional status. This raises the possibility that an apparent negative impact of nutritional status on educational outcomes may simply be due, at least in part, to the fact that households that live far from schools also live far from health clinics.

The need to collect data on all child, household, and school variables, particularly variables related to school (and teacher) characteristics, implies that a very large data collection exercise is needed. There are three exceptions to this general need for detailed school data. First, if the main objective of the household survey is to assess the current state of affairs and to examine correlation but not necessarily causation, it is only necessary to collect data on the variables of particular interest to the researcher, most of which can be collected at the household or child level. For example, if what is being studied is simply which socioeconomic groups have the lowest levels of school enrollment, there is no need to collect data from schools to measure this. Second, if school characteristics can be altered as part of a randomized evaluation, it is not necessary to collect data on all variables thought to be causal factors. (Randomized evaluations are discussed in detail at the end of this section.) Third, in nonrandomized settings there is one approach that can be used to estimate the impact of child and family characteristics on educational outcomes in the absence of data on school characteristics and prices. This approach will work for any household survey that uses a two-stage (or three-stage) sample design (which is by far the most common method for drawing a sample of households). Such a sample design makes it possible to use community or school "fixed effects" estimation techniques on a single cross-sectional data set. The idea is that all households in the community face the same school variables. This approach can be used both when a community has only one school and when there are

several schools in the community from which parents can choose.

If a linear model can be assumed (with no interaction effects between school characteristics and child or family characteristics), the combined effect of all school characteristics for each community can be captured by a dummy variable for that community. Variation of household and child characteristics within each community can be used to estimate the impact of these variables on educational outcomes. One disadvantage of this technique is that it cannot estimate the impact of school variables on educational outcomes.<sup>12</sup>

To summarize, in nonrandomized settings there are two ways to estimate the determinants of educational outcomes using a single cross-section of data. First, after collecting detailed information on child, household, and school characteristics, including school fees, for all local schools in each community, the researcher can do a "full" estimation of the determinants of educational outcomes. The second option is to dispense with the collection of school-level data and instead use a fixed effects procedure to estimate the impact of family- and child-specific variables (but not school variables) on educational outcomes. The remainder of this subsection discusses what can be learned about specific policy issues using cross-sectional data.

**ISSUES CONCERNING THE PRICE OF SCHOOLING.** What kind of cross-sectional data should be collected to examine the four questions on education pricing policies posed in the first section of this chapter? The first question is the impact of school fees on educational outcomes, which raises issues that apply to the other questions. The above discussion suggests that to estimate how mandatory school fees affect educational outcomes, it is necessary to collect data on both the quality and prices of local schools and then to estimate a model in which educational outcomes are determined by child and family characteristics, school quality, and school fees. If households have a choice of schools, it is still possible to estimate a similar model after correcting for selectivity bias due to school choice, as in Glewwe and Jacoby (1994). Reduced form relationships are the least difficult to estimate, but it may also be useful to estimate a production function (if the educational outcome of interest is learning) or a conditional demand relationship to probe deeper into the ways in which school prices determine educational outcomes.

It may be tempting to collect only price data from schools if the researcher is interested only in education pricing policies. However, price data are not sufficient for analyzing these issues. Data on many other school characteristics must also be collected to avoid omitted variable bias. Schools that charge high fees may well be high-quality schools, and looking only at prices could lead to the (mistaken) inference that high prices lead to high school enrollment rates because the (unmeasured) characteristics that make a school good persuade parents to enroll their children in that school. Most previous LSMS surveys, the exceptions being surveys in Ghana, Jamaica, Morocco and Vietnam, have not collected detailed data on school quality and school prices. Collecting these data is not a trivial task. Even with data on local school prices and school quality, the omission of some aspects of school quality that are difficult to observe and measure, such as teacher motivation, can lead to omitted variable bias.

Is there any way to examine the impact of school fees on educational outcomes without collecting detailed data on schools? Yes, if the researcher is willing to make some assumptions about travel time costs and the price of schooling. The idea is that each household in the community lives a different distance from the local school, and this variation in distance is equivalent to variation in tuition costs within each community. If this equivalence holds, it is possible to estimate how school fees affect educational outcomes.<sup>13</sup> Using data on hourly child wage rates, the amount of time it takes a child to walk to school can be transformed into a price that varies among children within the community due to variation in distance. Child wage rates can be obtained from either the community questionnaire or, in some cases, the labor section of the household questionnaire. Note, however, that the assumptions that must be made to use this method may be false, so there is some risk involved in using it. There is little empirical evidence on the extent to which estimates based on this method can be misleading. For examples of the application of this method see Gertler and Glewwe (1990, 1992) and Selden and Wasylenko (1995).

To summarize, there are two ways to estimate the impact of mandatory school fees on educational outcomes using cross-sectional data. First, detailed information can be collected on child and family characteristics and on both the characteristics and the fee structure of local schools, from which a reduced form

relationship of the impact of these variables on educational outcomes can be estimated. It may also be feasible to estimate some conditional demand functions. The second option is to collect data only on the distance or travel time between each household and the local school or schools, and use differences in travel time or distance to estimate a reduced form relationship. If the assumptions underlying these estimates are accurate, which is debatable, the estimates will measure how educational outcomes change when mandatory fees are changed.

Using distance within a community to understand the impact of prices on school attainment leads to the third pricing policy issue—the impact of school distance on educational outcomes. Taking the first route described in the previous paragraph—full estimation—the explanatory variables already include distance (or travel time), the various prices, and other school variables. It is then possible to estimate the impact (reduced form or otherwise) of all these explanatory variables. If it is reasonable to use variation in distance or travel time to estimate price effects, the impact of distance (or travel time) and mandatory school fees should be the same (after assigning a monetary value to children's time). Thus this method can be tested. Taking the second route (based on variation in distance within communities), the answer is already clear, but it is not possible to determine whether the price and distance effects are two distinct effects or two different ways of measuring the same effect.<sup>14</sup>

The second question posed in the first section was: what is the impact on educational outcomes of changes in prices for specific learning materials not provided by the school, such as textbooks, pencils, and extra classes? To answer this question it is necessary to regress the educational outcome being studied on child and family characteristics and the full set of school quality and school fee variables. In contrast to the situation concerning school fees and distance, there is no way to avoid collecting detailed school data to answer this question. The simplest approach is reduced form estimation, but again it may also be useful to try to estimate a conditional demand relationship. The school fee variables should include mandatory school fees and the prices of the specific inputs. This kind of estimation is rarely done because prices on specific educational inputs are seldom collected in household surveys—an oversight that should be corrected in future surveys.

The fourth question about pricing educational services involves the effects of introducing a student loan program, including who would receive the loans and what impact receiving a loan would have on the educational outcomes of the recipients. Student loan programs are fairly rare, so it is not possible to evaluate them in most countries. Even where such programs exist, in most cases they are national programs and thus do not vary across households at a single point in time. Therefore it is not possible to evaluate student loan programs using cross-sectional data.<sup>15</sup>

Before turning to the next set of issues, an important caveat must be recognized with regard to the use of cross-sectional data to analyze the impact of pricing policies on school outcomes. The methods discussed above for estimating the impacts of school prices on educational outcomes assume that none of the prices associated with schools can be changed unless the parents have a number of local schools from which to choose (the scenario in Figure 7.2). Yet this assumption might be challenged using the three potential problems discussed in the previous subsection—household migration for schooling purposes, household actions that change local school characteristics, and government consideration of local conditions when making decisions on local school characteristics and prices. In general, these problems cannot be resolved using cross-sectional data (the exception being the migration problem; see the discussion of panel data in Chapter 23). They will be discussed below in the subsection on panel data.

A final practical matter should be discussed regarding the collection of school price data as part of a household survey. While it is appropriate to collect data on schooling expenditures in the household questionnaire, it is important to collect information on mandatory fees and other prices directly from schools in either a community questionnaire or a school questionnaire. There are several reasons for this. First, it is likely that information gathered from households about the prices they paid will contain recall errors, and such measurement error can lead to biased estimates. Second, there is often more than one school available in a local community, and it is difficult to ascertain which prices reported by households apply to each school. Third, in poor communities it is possible that no one in the households surveyed is enrolled in secondary school, so no information on secondary school fees can be obtained from the household ques-

tionnaire. Of course, it is still important to collect expenditures on school fees in the household questionnaire to calculate total household expenditures and to check the accuracy of the data collected from the community or school questionnaire.

**ISSUES CONCERNING SCHOOL AND TEACHER CHARACTERISTICS.** Investigating the impact of school and teacher characteristics on educational outcomes requires detailed information on both school characteristics and school prices. There is no way to avoid collecting detailed data on local schools (unless one undertakes a randomized evaluation, a major activity that is discussed in detail below). Before turning to the four types of policy questions regarding school and teacher characteristics posed in the first section of this chapter, it is useful to consider whether to estimate a reduced form relationship, a conditional demand function, or a learning production function. Estimating a reduced form relationship amounts to regressing the educational outcome of interest on all exogenous determining factors, which consist of school characteristics, school prices, and the child and family characteristics that can reasonably be assumed to be exogenous. A full set of school quality variables should be used to minimize omitted variable bias. However, there are potentially serious problems regarding whether local school variables are endogenous. These problems will be discussed further in the panel data section below.

In some cases the researcher may be particularly interested in estimating a learning production function. This implies regressing skills learned on school attendance, school quality, years of schooling, child and family characteristics, and expenditures on “optional” items (such as textbooks, exercise books, and extra classes). School expenditures are clearly endogenous, so instrumental variables—variables that determine school expenditures but do not directly affect learning—need to be found; obvious candidates are the prices of the specific items. Daily attendance is also clearly endogenous, and an added difficulty is that attendance data over long periods of time are hard to obtain. Possible instruments for attendance are household wealth, distance to the school, and the productive assets relevant to self-employment. (If data on daily attendance are not available, production function estimates will suffer from omitted variable bias, so one must turn to either reduced form or conditional

demand estimation.) The number of years of schooling is clearly endogenous, but the number of years of schooling of the children who are currently in school may not be. In particular, if virtually all children in the survey begin their schooling at the “normal” age (usually, six years) and very few have already finished their schooling (if, for example, the estimates are for primary school students in a country where children rarely drop out while in primary school), years of schooling can be considered exogenous.<sup>16</sup>

In practice, it is difficult to collect information on daily attendance and, in some cases, on school expenditures. Thus it may be necessary to choose between reduced form and conditional demand relationships. For example, it would be very useful to know how textbooks affect learning, but in school systems in which parents purchase (or decide not to purchase) textbooks for their children, this cannot be estimated as a reduced form relationship because possession of a textbook is a decision made by parents. A reasonable instrumental variable, such as the price of textbooks, raises the possibility of estimating a conditional demand relationship even in the absence of data on school attendance, and perhaps even in the absence of data on expenditures on other school supplies. This can be done by replacing (“substituting out”) the variables for which there is no information (such as school attendance) or no particular interest with the exogenous variables that determine those variables, and then predicting (“instrumenting”) any remaining endogenous variables. This yields an estimate of the impact of textbooks on learning that is conditional on current household circumstances and behavioral responses to those circumstances. Although this is not a production function, if the estimates are accurate they can be used to predict the impact on learning if schools were to provide textbooks free of charge.

Unfortunately, there is a serious problem with conditional demand estimation. Replacing endogenous variables such as school attendance and school expenditures with the exogenous variables that determine them may use up all the instrumental variables that were available to use as instruments. For example, if the researcher wants to estimate the impact of textbooks but there are no data on school expenditures, the school expenditures variable will have to be “substituted out.” Total school expenditures probably depend on the price of textbooks. Thus substituting out that variable will introduce the price of textbooks

into the equation to be estimated, which means that textbook prices cannot be used as an instrumental variable for textbooks. Overall, economists do not agree on whether conditional demand estimation is likely to be feasible in general, and each case must be judged on its own specific circumstances.

Returning to the distinction between the four types of school characteristics, it was pointed out in the first section that, in theory, school management practices and school admissions and advancement policies have only indirect effects since their only effect is to change what happens in the classroom—by changing material inputs, teacher characteristics, and pedagogical techniques—and to regulate who can enter the classroom. Thus, if data are available for all of the important material input and teacher variables, only the material input and teacher variables are needed in any regression estimated to assess the impact of these school quality variables on educational outcomes. This implies that it is not necessary to collect data on school management practices or admissions and advancement policies if these aspects of schooling are not being studied.

However, there are three reasons why data on school management, admissions, and promotion policies should be collected (and why the school questionnaire introduced in the third section of this chapter does collect them). First, it is probably impossible to collect data on all material input and teacher variables (indeed, it is hard to imagine how data could be collected on a variable such as teacher motivation), and school management variables may pick up some of the effects of these variables—reducing omitted variable bias.<sup>17</sup> Second, many policymakers and researchers are very likely to be interested in these issues, either now or in the future. Third, there is very little cost to collecting data on school management practices and admissions/advancement policies if the survey has already been designed to collect data on material inputs, teacher characteristics, and pedagogical techniques. All that needs to be done is to add a few questions to the school questionnaire.

Finally, it is possible to combine estimates of price effects with estimates of school and teacher characteristic effects. This is useful for assessing the feasibility of simultaneously raising school fees and using the funds to pay for improvements in school quality. “Full” reduced form estimates of educational outcomes will use both price and school quality variables. This will

provide parameter estimates that can be used to predict the net effect on those outcomes of changing prices and school quality (more specifically, changes in material inputs and teacher variables) simultaneously. In many cases, it is useful to interact the school and price variables with household income in order to see whether the effects depend on household income levels. For examples and discussion of this approach see Gertler and Glewwe (1990, 1992) and Selden and Wasylenko (1995).

In summary, analyzing the impact of school and teacher variables on educational outcomes requires detailed information on school (and teacher) characteristics and school prices. It is impossible to avoid collecting detailed data from local schools (unless one implements a randomized evaluation). Information should be collected both on what takes place inside the classroom and on school management and admissions/advancement policies.

**OTHER POLICIES THAT AFFECT EDUCATIONAL OUTCOMES.** The first section of this chapter listed three other government policy areas that can affect educational outcomes in developing countries: policies related to child health and nutrition, policies that affect the opportunity cost of a child's time, and policies that affect the returns to schooling in the labor market. It is not clear that a single cross-sectional household survey can shed any light on the last type of policy. This is because a single household survey identifies relationships based on variations across households and communities at one point in time. Most government policies that affect labor market outcomes (for example, a freeze on hiring new government workers) usually do not vary much across regions and communities at a single point in time. However, some policies may pertain to specific regions or zones. For example, if a country declares some areas to be export promotion zones and if employment opportunities in these zones generate a higher payoff for schooling, it may be possible to estimate the impact of the existence of these zones on the demand for schooling and thus to investigate what would happen to educational outcomes if these export promotion zones were extended into other areas. However, this kind of estimation has few implications for designing the education module or any other part of a multitopic household questionnaire. All the researcher needs to know is the location of the communities sampled in the survey.

Regarding policies that affect the value of children's time, community questionnaires in multitopic household surveys should collect data on child wage rates. The community questionnaire introduced in Chapter 13 does collect such data. A problem often encountered when collecting such data is that some communities have no wage labor market for children. However, in most communities and in almost all rural areas, many children work on the family farm or for the family business. It may be possible to estimate the value of children's time by estimating a profit function for a farm or business—for example, using the approach of Jacoby (1993) or Newman and Gertler (1994). Such estimates are particularly useful because they are based on variation within the community, so they can be estimated even when using a community fixed effects estimation procedure to control for unobserved school and community characteristics. For the data needed to estimate profit functions for farms or nonfarm household enterprises, see Chapters 18 and 19.

Another way in which government policies can affect the opportunity cost of children's time is by altering the amount of time children spend doing household chores—say, by providing a well (which would reduce the time that children, especially girls, spend fetching water) or changing the prices of kerosene and cooking fuel (which would affect the time that children spend fetching firewood). In general, this type of information should be collected in a community or price questionnaire (see Chapter 13). The empirical techniques are little different from those already discussed above. However, if a community fixed effects estimation procedure is used to control for unobserved community variables (such as school quality), it is not possible to address these kinds of questions unless there is variation in the relevant distances (for example, the distance from each household to a well) within a community.

Finally, government policies that affect child health and nutritional status can also affect educational outcomes. Clearly, data on health and anthropometric status must be collected in the respective modules of the household questionnaire (see Chapters 8 and 10 for details). Unlike many other child characteristics (such as sex, innate ability, and birth order), child health is not exogenous. Parents and their children make choices that affect the children's health. Thus estimates of the impact of child health and nutrition on educational outcomes must be either conditional demand relation-

ships or learning production functions, and the kind of instrumental variables needed are those that predict child health but that have no direct impact on educational outcomes.<sup>18</sup> Some obvious instruments are the availability, quality, and fee structure of local health clinics and the prices of common medicines. However, if community fixed effects are used to control for unobserved school quality, these variables cannot be used as instruments. The exceptions are distance or travel time, which require data from each household on the distance to the local health facilities. Alternatively, it is possible to use variations in parental height. For a fuller discussion of what predicts child health and anthropometric status, see Chapters 8 and 10.

A final word of caution is in order regarding child health and schooling. Recent research has seriously questioned whether cross-sectional data can be used to estimate the impact of child health on educational outcomes. In particular, it is hard to find instrumental variables that predict child health status but do not also directly affect educational outcomes. For example, the prices of medicines should, like all other prices, directly affect parental decisions regarding child schooling. Many researchers have tried to overcome these problems by using panel data. This method will be discussed in the next subsection.

**SUMMARY.** The main conclusions of this discussion have been as follows.

- If the researcher is interested only in understanding the impact of educational outcomes on other socioeconomic outcomes, there is no need to collect data from local schools, since such data are useful only for understanding the impact of school policies on educational outcomes.
- If the researcher is interested only in the impact of child and household factors on educational outcomes—and not in the impact of school or community variables on educational outcomes—there is no need to collect detailed data on schools.
- If the researcher is interested only in the impact of mandatory school fees or distance to schools on educational outcomes, this relationship can be estimated without collecting detailed school data, by collecting data on the distance from each household to local schools (although this requires making certain assumptions that may be false).
- It is necessary to collect detailed data on local schools in order to examine the impact of other

pricing policies and of classroom and teacher characteristics (including pedagogical practices) on educational outcomes.

- In examining school fees it is important not to depend only on the expenditure data on education collected in the household questionnaire; information on school fees must also be collected from a school or community questionnaire.
- In general, it is not possible to analyze the impact of student loan programs on educational outcomes using cross-sectional data, even if detailed data are collected on local schools.
- It is very difficult, if not impossible, to estimate a production function for learning using cross-sectional data, and some would argue that the same is true for estimating conditional demand relationships.
- As long as a household survey is designed to collect detailed data on what goes on in the classroom (material inputs, teacher characteristics, and pedagogical practices), it should also collect information on management and admissions/advancement policies, even if policymakers at the time are not particularly interested in analyzing how school management and admissions/advancement policies affect educational outcomes.
- A single cross-sectional data set cannot shed light on the impact of government labor market policies on educational outcomes unless the policies vary by region; however, a series of cross-sectional data sets over a period of time in which those policies change may be able to do so.
- Information should be collected on the education levels of the parents of all household members, regardless of whether the parents live in the household or are even alive.

#### *Using Panel Data to Investigate the Determinants of Educational Outcomes*

Additional estimation techniques can be used if panel data have been collected—in other words, if the same households have been interviewed at more than one point in time over a period of months or years. Panel data are relatively rare in developing countries, but this need not be the case in the future.

Many problems arise when basic multiple regression techniques, such as ordinary least squares, are used on cross-sectional data because some of the variables included in the regression are correlated with variables that are not included. The most obvious way to solve

this problem is to collect information on the missing variables, which will reduce this omitted variable bias. In theory this should be effective, but it is rarely possible to collect all possible variables, either because of cost limitations or because some variables (such as parents' tastes, children's innate ability and teachers' motivation) are inherently difficult to measure.

In principle, instrumental variable methods, such as two-stage least squares, can overcome the problem of an explanatory variable being correlated with unobserved terms. Instrumental variable techniques can also resolve many problems of measurement error. Unfortunately, finding credible instrumental variables is not easy. (See Strauss and Thomas 1995 for a detailed discussion.) Another approach is to use fixed effects procedures to control for differences across communities or, in some cases, across families. Yet community fixed effects procedures cannot eliminate problems due to variation in unobserved variables within the community, and even family fixed effects cannot deal with variation within the family (such as variation in the innate abilities of siblings). Some of these problems can, under certain assumptions, be circumvented by using panel data. Chapter 23 provides a general overview of the advantages and problems associated with the collection and analysis of panel data. The current subsection discusses four ways in which panel data can be used to examine education issues in developing countries.

**VALUE-ADDED SPECIFICATION OF THE DETERMINANTS OF LEARNING.** Estimating the determinants of learning (whether production functions, reduced forms, or conditional demand relationships) is of great interest to policymakers. Many problems arise when trying to estimate these determinants using cross-sectional data. One problem is that learning is a cumulative process in that what a child has learned up to a particular point in time depends on an entire "history" of school, household, and child variables. Collecting data on past events and conditions is difficult, but omitting these variables may result in serious problems of omitted variable bias. Specifying an equation of the determinants of learning in "value-added" form can alleviate some of these problems. To see this, consider a simple linear model of learning in, say, the first three years of school:

$$(7.1) \quad A_1 = \beta_0 + \beta_1 X_1$$

$$\begin{aligned} A_2 &= \beta_0 + \beta_1 X_1 + \beta_2 X_2 \\ A_3 &= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \end{aligned}$$

where  $A_1$ ,  $A_2$ , and  $A_3$  represent achievement (for example, test scores) at the end of grades 1, 2, and 3, respectively.<sup>19</sup> The constant term  $\beta_0$  represents what, if anything, the child has learned before entering first grade. The  $X_1$ ,  $X_2$ , and  $X_3$  variables refer to the values of all child, school, and household variables in grades 1, 2, and 3, respectively.

The relationships shown in equation 7.1 imply the following relationships:

$$(7.2) \quad \begin{aligned} A_2 - A_1 &= \beta_2 X_2 \\ A_3 - A_2 &= \beta_3 X_3 \end{aligned}$$

Estimation of the relationships in equation 7.2 implies much simpler data collection. For example, to understand the effects of the child, household, and school characteristics that prevail when children are in third grade (in other words, to estimate  $\beta_3$ ), the  $X_3$  variables can be regressed on  $A_3 - A_2$ . No information is needed on  $X_1$  or  $X_2$ . More generally, if data are available on test scores from a particular past period in time, there is no need to collect any data for any previous time periods. Estimating equation 7.2 requires panel data because it requires that each child be tested at two different points in time.<sup>20</sup>

Finally, some economists have suggested that using panel data to estimate a value-added production function has the added advantage of eliminating all unobserved fixed effects (Hanushek 1992). This is true under certain functional form assumptions, but it is somewhat risky because it requires the assumption that the unobserved variables affect total learning but do not affect increments to learning, which seems implausible.

**IMPACT OF EARLY CHILDHOOD NUTRITION ON SCHOOLING.** It is widely thought that poor nutrition in childhood leads to poor school performance (Del Rosso and Marek 1996; Pollitt 1990). This implies that policies improving children's nutritional status may also improve their school performance. However, using cross-sectional data to estimate the impact of childhood nutrition on educational outcomes is difficult or even impossible, for two main reasons. First, the crucial factor governing a child's nutritional status may be the conditions that prevail in the child's household

during his or her first two to three years of life rather than the conditions at the time the child is in school. Cross-sectional data can provide only a summary measure of nutritional status over a child's life—say, height for age—and this summary measure is likely to contain substantial measurement error. Second, children's nutritional status is clearly endogenous, which means that instrumental variables would be needed in order to obtain unbiased estimates. Finding credible instrumental variables is not an easy task.

Panel data can overcome both of these problems. By definition, they can provide information on nutritional status in the first years of life and on subsequent school performance for the same child. In addition, panel data can provide a rich source of instrumental variables that are more plausible than those based on cross-sectional data; such instrumental variables include price shocks (Alderman and others 1997) and indicators of health during early childhood (Glewwe, Jacoby, and King Forthcoming). More generally, because a child's nutritional status is determined in part in the first two to three years of life—before he or she starts attending school—the process by which child nutrition affects school performance is inherently dynamic, which leads to the need for panel data (see Chapter 23 for further discussion of this point).

**RETENTION OF SKILLS LEARNED.** One issue that has received little attention in the past is the retention of skills acquired in school. Some educators claim that a minimal number of years of school attendance, such as four or six, is needed for children to avoid lapsing into complete illiteracy after leaving school. In addition, the jobs young people take after leaving school will result in loss, retention, or increase of the skills learned in school, depending on the type of job. Exactly what happens to skills after students leave school is unclear and may have important policy implications. For example, if a certain number of years of schooling guarantees lifelong literacy for nearly all those who leave school, a reallocation of educational resources may be needed to ensure that all children acquire this minimal level of schooling.

To analyze students' skill retention after leaving school, it is necessary to measure the skills of the same individuals at two or more points in time—at the time they leave school and at one or more later points in time. Unfortunately, few examples of such data exist because measuring cognitive skills after students leave

school cannot be done in the classroom. Thus panel data are needed.

**ENDOGENOUS SCHOOL CHARACTERISTICS.** As was pointed out above, school characteristics can be endogenous either because of selective migration or because households' characteristics or actions determine local school characteristics. In general, as long as some cross-sectional data already exist on migration, selective migration does not necessarily imply that panel data must be collected. In contrast, if local school characteristics are determined in part by unobservable community characteristics, problems may arise that would be difficult to overcome without panel data. Both of these points are explained in the discussion on evaluating social sector programs in Chapter 23.

Only one empirical study on educational outcomes has used panel data to examine the impact of changes in local school characteristics on changes in those outcomes: an analysis of Indonesian data by Pitt, Rosenzweig, and Gibbons (1993). Several comments are in order regarding this study. First, the authors were able to implement their methodology using a district-level panel data set constructed from a series of cross-sectional household surveys; thus they did not need panel data on households. Second, the data sets used contained well over 100,000 households—many more than in even the largest LSMS survey. Third, in principle, such panel data are not needed if retrospective data exist on the availability of schools. A final, more general, comment is that methods to avoid bias brought about by endogenous school characteristics are still quite experimental; more research is needed on how to reduce such bias.

While these four examples show how panel data can be useful in analyzing education issues in developing countries, panel data may also have some disadvantages. First, because panel data tend to magnify the bias brought about by measurement error, careful thought must be given to which instrumental variable techniques to use—which may dictate that the panel should extend over three time periods—when making plans to collect panel data. Second, collecting panel data can lead to serious problems of sample attrition. Third, panel data may be more expensive to collect, relative to a series of cross-sectional surveys. On the other hand, the expense will be relatively low if individuals or households that move are not followed; see Chapter 23 for a detailed discussion of the costs of collecting panel data under different scenarios.

### *Using Randomized Trials for Program Evaluation*

For researchers interested in collecting data that make it possible to assess how government policies affect educational outcomes, the discussion so far may have been discouraging. Many problems can arise regardless of whether cross-sectional or panel data are used, which raises serious questions about whether it is possible to estimate the effects of changes in educational policies with any accuracy. Economists have become more aware of these problems in recent years, and some now argue that randomized trials offer the best hope for overcoming these problems (Newman, Rawlings, and Gertler 1994; Burtless 1995).

The concept behind randomized trials is quite simple and persuasive. If a researcher wants to examine the impact of a particular educational policy, he or she should choose a representative sample of schools or communities and randomly divide the sample into two groups. The education policy of interest will be implemented in one group (the “treatment” group) while the other group will serve as a comparison group (the “control” group). This approach is theoretically sound, but there are some problems with putting it into practice.

First, governments are often reluctant to allow randomized trials because: they may be unwilling to admit that they do not know what works best; they are not patient enough to wait for the results; the results may go against entrenched political interests; they do not want to be seen as making their citizens subjects in “human experiments”; and they do not like that fact that random trials inevitably involve denying assistance, at least initially, to some people who are thought to be particularly needy. A second problem is that it may be difficult to prevent some of the people assigned to control groups from participating in the “treatment” group. For example, in a recent study of primary education in Kenya (Kremer and others 1997), the schools that received assistance experienced large increases in student enrollment, especially in the lower grades—which may have “contaminated” the experiment. (The problem of getting a clean comparison is highlighted by Heckman and Smith 1995.)

A third difficulty is the ethical issues that some people claim are involved in using human beings as participants in experiments. However, many things can be done to overcome such objections; indeed, medical researchers have been conducting randomized trials

on human beings for decades and have developed rigorous procedures that meet contemporary ethical standards. Finally, randomized trials can be costly. This is mainly a problem if the trials cannot be carried out as part of the implementation of a policy or project, which forces the program being evaluated to be financed out of scarce research funds. However, in many cases randomized trials can be designed as part of an existing government project (Dow and others 1997), reducing costs substantially. Working with non-governmental organizations is another way to reduce costs by evaluating existing projects; this was done by Glewwe, Kremer, and Moulin (1999).

A final issue to consider when conducting randomized trials is whether a multitopic household survey is needed. Randomized experiments can be designed to evaluate specific educational policies without conducting a household survey. In particular, data need be collected only on the socioeconomic outcomes that the policy is expected to influence. For example, it may not be necessary to collect information on household consumption expenditures because the policy was not expected to change them. However, there are three reasons why a multitopic survey should be implemented in the context of a randomized evaluation. First, it may be useful to perform separate comparisons for different socioeconomic groups, such as poor and nonpoor households. Second, it may also be useful to compare results from a randomized evaluation with results from cross-sectional or panel surveys, which implies the need to collect additional data as explained in the previous subsections. Third, the policy being examined may affect more socioeconomic outcomes than initially expected.

This being said, it may not be necessary to organize a nationwide multitopic household survey to accompany a randomized evaluation. In practice, randomized evaluations often cover a relatively small geographic region in order to contain costs. If the evaluation is not designed to be nationally representative, there is no need for the accompanying survey to be nationally representative either. On the other hand, randomized evaluations can sometimes be performed for a large geographic region or even a nationwide sample. One example of this is the random phasing-in of a new pricing policy over several years, as was done for health care services in Indonesia (Dow and others 1997). In such cases a household survey that covers several provinces or even an entire

country may be the best way to assess the impact of the new policy or program.

### **Proposed Draft Education Modules for LSMS Surveys**

This section introduces three draft education modules that can be used in future multitopic household surveys: a short version, a standard version, and an expanded version. (The modules themselves are provided in Volume 3.) The section also introduces two draft questionnaires for collecting data from local schools. Each draft module or questionnaire should be tailored to meet the overall objectives of the survey and the characteristics of the education system in the country where the survey is being conducted. When designing household questionnaires, survey designers may want to incorporate features from more than one of the three draft modules. For example, they may choose to use the short module as the basis for the survey but add several questions from the standard module. Ultimately, the objectives and constraints will be different in each country, so these modules should be thought of as starting points rather than finished products.

Box 7.2 summarizes the capacity of each of the different modules to answer the policy questions raised in the first section of this chapter. The box does not address the short education module, because this module is designed primarily for investigating the impact of household members' educational outcomes (grade attainment, years of schooling, repetition, and diplomas obtained) on other socioeconomic outcomes. Other chapters in this book examine in detail how to conduct such investigations.

Finally, as was mentioned several times above, certain data from other parts of a typical multitopic household survey are very important for analyzing the determinants of educational outcomes. These data are:

- Child wage rates (from the community questionnaire).
- Schooling and occupation of the parents of all household members (from the household roster).
- Prices of optional educational inputs—textbooks, exercise books, slates, pencils—from the price questionnaire.
- Anthropometric measurements and other indicators of child health.

- Distances and travel times to community sources of water and fuel (to calculate the opportunity cost of children's time).

### **Comments and Notes on the Draft Education Modules**

This section provides detailed comments and notes that explain why the modules introduced in the previous section (and presented in Volume 3) take the form that they now have. The first three subsections discuss each household-level module in detail, including how each can be altered to fit specific needs and constraints. The fourth subsection discusses the collection of school data using school and teacher questionnaires.

#### *Short Education Module*

The short module is designed for a survey that focuses on a topic other than education and for which the standard education module is too large. The short module collects data on the most basic educational outcomes: years of schooling completed, degrees or diplomas obtained, current school attendance, repetition, and expenditures on schooling. This module is appropriate if policymakers and researchers have little or no interest in estimating the determinants of educational outcomes, in which case education data are needed primarily to provide basic descriptive statistics such as current school enrollment rates and the distribution of education across the adult population, as well as estimates of the impact on other socioeconomic outcomes of years of schooling, repetition, and certificates and degrees obtained. In such situations there is no reason to collect data on education from either the community questionnaire or a school questionnaire.

The following notes explain several details of the short module

**Q3–4, Q6–7.** These questions are almost identical, except that Questions 3 and 4 are asked of people who have finished their schooling while Questions 6 and 7 are asked of people currently in school. This distinction (made in Question 2) is important to ensure that the response from individuals currently in school is their current grade rather than the “highest grade completed,” which, precisely interpreted, would be the grade immediately preceding their current grade.

**Box 7.2 Policy Issues and How the Draft Modules Can Address Them***Standard Module without School Questionnaire*

- Impact of school fees on all educational outcomes except learning (assuming that distance to schools affects educational choices in the same way that money costs of schooling affect those choices)
- Impact of school management on all educational outcomes except learning (assuming differences in school management can be summarized using a simple variable—such as public/private status of schools—that can be collected in the education module)

*Standard Module with School Questionnaire*

- Impact of school fees on all educational outcomes except learning
- Impact of the prices of optional schooling inputs on all educational outcomes except learning
- Impact of distance on all educational outcomes except learning
- Impact of material inputs and very basic teacher characteristics on all educational outcomes except learning, assuming (unobserved) teaching practices and more detailed teacher characteristics are only weakly correlated with (observed) material inputs and basic teacher characteristics
- Impact of government policies that affect the value of children's time on all educational outcomes except learning

*Expanded Module (Including Cognitive Tests, School Questionnaires, and Teacher Questionnaires)*

- Impact of school fees on all educational outcomes
- Impact of optional schooling input prices on all educational outcomes
- Impact of distance on all educational outcomes
- Impact of material inputs, teacher characteristics, and teaching practices on all educational outcomes
- Impact of school management on all educational outcomes
- Impact of school admissions and advancement policies on all educational outcomes (assuming that these policies vary across schools)

Source: Author's summary.

- Impact of government policies that affect the value of children's time on all educational outcomes

*Issues and Methodologies That Require Panel Data*

- Impact of student loan program—if country recently introduced a student loan program, and the collection of panel data began before the program was introduced and continued after it was introduced
- Impact of school characteristics on educational outcomes—if endogenous program placement biases conventional estimates
- Impact of school admissions and advancement policies on educational outcomes—if the policies changed recently and panel data were collected both before and after the policy change
- Impact of labor market policies on educational outcomes—if policies have changed in recent years and panel data were collected both before and after the policy change
- Impact of child health and nutrition on educational outcomes
- Estimated value added specification of learning production function
- Estimated retention of cognitive skills after children leave school

*Issues That Cannot Be Addressed Using Household Survey Data (Except through a Randomized Evaluation)*

- Impact of student loan programs (if country has never had a student loan program)
- Impact of school admissions and advancement policies on educational outcomes if the policies have not changed recently or if they have changed but panel data were not collected both before and after the change
- Impact of labor market policies on educational outcomes if policies have not changed in recent years or if panel data are not available for both before and after a policy change

**Q4, Q7.** Most diplomas are sequential, so that if a respondent has a higher diploma (for example, a diploma certifying completion of upper secondary school), this implies that he or she has also attained a lower diploma (for example, a diploma certifying completion of lower secondary school). However, there may be some ambiguous cases. If such cases are common, it may be that two or even three responses should be allowed for Questions 4 and 7.

**Q8.** The main purpose of this question is to see who really benefits from public education. For example, if most children from wealthy households go to private schools, the provision of public schooling may disproportionately benefit poorer households. Some policymakers and researchers may also be interested in seeing whether the impact of education on some other socioeconomic outcome, such as wages, also varies across public and private schools. If this is of interest, a

**Box 7.3 Cautionary Advice**

- *How Much of the Draft Module Is New and Unproven?* The short version of the education module follows a similar approach to past practice in LSMS surveys, and thus is based on proven methods. In the proposed standard education module, a few elements are new or have been used in only two or three past LSMS surveys, but none of these innovations is very complicated. Asking respondents to read a short sentence and perform simple written mathematical exercises has been tried in Morocco, South Africa and Vietnam. Asking separate sets of questions for persons now in school and for persons no longer in school (Question 6 of Part A) and asking separate questions for different levels of school (Question 7 of Part A) have been tried twice before (Jamaica 1990, Vietnam 1997–98). The questions on grade repetition (Questions 41–46 of Part A) have been tried in about four previous LSMS surveys. Part D on distances to local schools has been tried in two previous LSMS surveys (Ghana 1988–89, Vietnam 1997–98).
- *How Well Has the Module Worked in the Past?* In general, the education module has worked quite well. The only problems in past LSMS surveys have been in expanded modules, particularly ones that have administered achievement tests to household members and ones that have tried to match household members to specific schools. Much was learned from these exercises, and those lessons have been incorporated into the advice given in this chapter.
- *Which Parts of the Module Most Need to Be Customized?* Several points need to be watched closely. First, a clear distinction between training and apprenticeship needs to be made. This definition may vary across countries. Similarly, a clear distinction needs to be made between general education and technical or professional training. Second, when defining schooling levels and diploma codes, clear instructions are needed on how “old” systems of education, which will be reported by older individuals in the sample, are coded. Third, in many countries there are kinds of school expenditures that are rarely found in other countries. Unusual categories in any given country can be obtained by asking officials from the ministry of education and participating in the pilot test of the questionnaire.

question similar to Question 8 should be inserted after Question 4; this new question would refer to the last school attended by the respondent.

**Q9.** The main purpose of this question is to obtain accurate data on school expenditures for the purposes

of calculating total household expenditures. A simple question in the household consumption module is likely to result in serious underestimation of actual spending. The exact categories of expenditure must be adapted to fit the country in question; these should be carefully checked during the pilot test of the questionnaire.

**Q10–11.** These questions on repetition can be dropped if repetition is relatively rare. However, when repetition is not rare, it is useful to distinguish between years in school and grades attained. Simple estimates of rates of return to a year of education may be seriously overestimated if they do not account for repetition (Behrman and Deolalikar 1991).

**Standard Education Module**

The standard education module is designed to gather the information needed to answer most of the policy questions discussed in this chapter. It is therefore designed to allow researchers to investigate not only the impact of education on other socioeconomic outcomes but also the determinants of most educational outcomes. The main omission from this module is that no effort is made to collect detailed information on cognitive skills. Collecting cognitive skills data will be discussed further in the following subsection.

Before discussing this module in detail, several general points need to be made. First, a cursory glance suggests that this module is substantially longer than the standard education modules used in full LSMS surveys in the late 1980s and early 1990s. This appearance is deceptive. The main reason there are more questions is that people who have already finished their schooling answer a set of questions that is different from the questions answered by people currently in school. The former answer Questions 7–24 of Part A while the latter answer Questions 25–40. Having these two sets of people answer different questions should make the interview go easier; in particular, it should avoid any confusion about whether the grade indicated for a person still in school is the current grade or the “highest grade completed,” which, technically speaking, is the previous grade. On the other hand, there are a few additional questions on repetition, and Part D is completely new. However, some questions have also been deleted, and Part D has to be filled out only once for the entire household rather than once for each individual (and households that

contain no school-age children do not fill it out at all). Overall, there is a small increase in the number of questions asked but this increase is not nearly as great as initial appearances may suggest.<sup>21</sup>

Another general point is that one may want to add a few qualitative questions asking households why they made the schooling decisions they did. While some might argue that these explanations are rationalizations and thus are not very accurate, others may find such information interesting, if for no other reason than to see what households claim are the reasons for the choices they have made. In particular, a question could be inserted after Question 5 in Part A asking: “Why have you never attended school?” Possible reasons are the same ones given for nonattendance in the education module of the community questionnaire (see Chapter 13). It may only be useful to ask this question for relatively young individuals. A similar question could be added after Question 6 of Part A: “Why are you no longer attending school?” Again, this may only be appropriate for younger individuals. A third place to ask a qualitative question is immediately after Question 30 of Part A; for persons who were absent from school in some of the past seven days one could ask: “Why were you absent from school during the past seven days?”

A third general point is that any survey that uses the standard version of the education module should also collect anthropometric data, since there is strong evidence that early childhood nutrition affects educational outcomes.<sup>22</sup> The collection of anthropometric data is discussed in Chapter 10. A final point is that two school questionnaires—the questionnaire for administrators and the questionnaire for teachers—could be used with the standard module. This is not necessary, but it will increase substantially the range of issues that can be analyzed (as seen in Box 7.2).

The following notes on specific questions in the standard module are useful for understanding the module’s design and thinking about how to modify the module to fit particular interests and constraints.

**A1.** In general, it is strongly advisable to interview a subject directly, with the possible exception of children age 10 and under, because asking someone else greatly increases the chance of errors and missing data. By indicating whether a person was interviewed directly, this question provides information on the accuracy of the data.

**A3–A4.** For Question A3, a simple sentence of six to eight words in one or more languages should be prepared based on the primary-school curriculum. It is best to have three or four variants to prevent some people in a household from doing well because they overheard someone else read the same sentence. If the language used by the household members is not the official national language, A3 should be applied twice, once for the national language and once for the language used by the household. Mathematics problems should be simple addition or subtraction at a level consistent with two to three years of primary education. In some countries (Bolivia is one example) students receive scores on national examinations each year, and this information is commonly kept by the household. In such cases, for children currently in school, Questions A3 and A4 should be replaced with the scores on these examinations.

**A8, A26, A49–A50.** These questions on preschool can be dropped if preschool education is rare. It is important to clarify the difference between preschool and daycare; most surveys only investigate preschool.

**A12, A34.** If there are 10 or fewer postsecondary institutions in the country, it is useful to add a question requesting the institution’s name and to include in that question explicit codes for each institution.

**A13, A17, A21.** It is important to distinguish between finishing a grade and passing an examination on the one hand and finishing a grade and not passing the examination on the other. A student who finished a grade but did not pass an examination should still be classified as having finished that grade.

**A15, A19, A23.** These questions on when schooling is finished are important for two reasons. First, if a person dropped out after only one or two months, it is important not to credit him or her with an entire year of schooling. Second, combining this information with the information on the time when the respondent first entered school (Questions A47 and A48) makes it possible to verify grade repetition using a computerized data entry package. (See Grosh and Munoz 1996 for a detailed discussion of computerization of data entry.)

**A17–A18, A21–A22.** If no diploma is associated with lower secondary or primary education, the interviewer can simply ask the second question of each pair.

**A27.** It is important to know whether a student lives away from home while attending school because this implies a large opportunity cost to the household in terms of the time that the student could have been spent working for the household after school or on weekends.

**A28.** Finding out the name of the school is important for matching data from the school or community questionnaire. Codes should be assigned as soon as possible, preferably before any household interviews have started (so that interviewers will already have a list of codes when filling out the questionnaire) and certainly before the team leaves the community. If official school code numbers are used by the ministry of education, these codes should also be added in a separate column. It is not advisable to depend on national codes alone because new schools and private schools often do not have official code numbers.

**A29–A30.** Ideally, it would be useful to know about all absences during the entire school year and even in past years. Absences during the previous week are a very rough indicator of the necessary information, but it is unlikely that children or their parents will be able to accurately recall absences over a longer period of time. In some countries schools may keep records, at least for the current year. If so, serious thought should be given to trying to obtain that information. This would involve a significant amount of work and would require pretesting. Finally, combining information on absences during the previous seven days with information on class time as given in the school questionnaire allows researchers to calculate time spent in school in the previous week. However, if it is common for children to attend school for only part of the school day (because they arrive late or leave early), a separate question should be asked regarding the number of hours the student missed in the previous seven days due to arriving late or leaving early.

**A41–A46.** These questions could be condensed or eliminated for any level of schooling in which repetition is rare—in other words, if fewer than 2–3 percent of students repeat. Condensing these questions simply involves asking whether the individual repeated any grade in primary or secondary school and, if so, which grades were repeated (allowing for up to three responses).

**A47–A48.** To verify the accuracy of the information, it is important to ask both questions.

**B1–B2.** These questions include professional and technical training. An alternative is to exclude such activities from the information collected here and to add to Part C a question on expenditures on professional and technical training during the past 12 months.

**B2.** This question has two purposes: accurately collecting total educational expenditures by a household and collecting child-specific information on how household resources are allocated across siblings and across different types of educational expenses.

**B3–B5.** This information on assistance received from people who are members of other households should be coordinated with information from the transfers and other nonlabor income module on income transfers received from other households (see Chapter 11). Rules need to be set to avoid double counting income that goes directly to the household. These rules may vary according to country-specific characteristics. Payments made directly to schools, such as tuition payments, should probably be counted only here, and not in the transfers and other nonlabor income module.

**B6–B7.** If a voucher scheme of some kind exists, separate questions along the lines of Questions B6 and B7 should be added. If there are different kinds of vouchers, another question should be added to determine which kind the child receives.

**C1–C3.** The definition of an apprenticeship needs to be carefully worded in order to reflect the system of apprenticeship that prevails in the country where the survey is being carried out. If apprenticeships are quite rare in the country—involving less than 2–3 percent of the population—Questions C1, C2, and C3 can be dropped.

**C1, C4.** It may be easier to obtain accurate responses to these questions by dividing each into two distinct questions. The first question would ask whether the respondent has ever been an apprentice (or had technical training) and the second would ask whether the respondent is currently an apprentice (or currently being trained). The best wording for these questions should be checked during the pilot test of the questionnaire.

**C4–C6.** It is important to distinguish between professional and technical training at the postsecondary level and equivalent training that does not require the student to have finished upper secondary school. Only the second kind of training is considered here; the first kind is addressed in Part A as general postsecondary education. Particular attention should be given to teacher training. In some countries teacher training always takes the form of postsecondary education, and thus should be dealt with explicitly in Part A. In other countries at least some types of teacher training are not considered postsecondary; these types should be explicitly identified in Part C.

**PART D.** These questions are intended to obtain a picture of what choices parents have when they send (or do not send) their children to school. When combined with information on the date when a school opened, these questions can be used to get around the problem of nonrandom placement of schools by the government. The ID codes are needed to match this information with the school that is chosen (obtained in Question A28) and with the detailed information obtained in the community questions or the school questionnaire. Since respondents may not be adept at providing distance information, this information can be augmented with data from GPS (global positioning system) measurements of the location of the households and the schools.

#### *Expanded Education Module*

The education module in the household questionnaire can be expanded in two general ways. First, relatively simple achievement tests can be administered to household members, both members currently in school and members out of school. This is a major undertaking, but one that when done correctly can yield rich data that will deepen policymakers' understanding of what makes schools perform effectively and how skills learned in school contribute to a wide variety of socioeconomic outcomes. (If this route is taken, Questions A2 and A3 in the standard module can be dropped.) Second, more information can be collected on children currently in school—which amounts to adding questions to Part B of the standard module. In this subsection some practical tips on administering tests are discussed, and some questions are presented that can be added to Part B of the standard education module.

**ACHIEVEMENT TESTS.** There is not enough space in this chapter to discuss how to design and administer achievement tests as part of a multitopic household survey. The next few paragraphs will discuss major testing issues and the advantages and disadvantages of the different testing choices. It is strongly recommended that a highly qualified local or international expert be hired to develop the tests. (For a good discussion of literacy tests in the context of household surveys see UNNHSCP 1989 and Wagner 1993.) The basic decisions that must be made are: who is tested; which skills are tested; the length of the tests; the relationship between the tests and the national curriculum; and where the tests take place.

There are essentially two types of household members that can be tested: children of school age and adults who have finished their schooling. Children are tested to find out how their characteristics, along with household and school characteristics, affect the skills that they acquire. Adults are tested to study the impact of skills learned on socioeconomic outcomes. It is often useful to test both groups in order to study the impact of determinants of learning among children on their socioeconomic outcomes in adulthood. For example, if a certain schooling improvement costs \$20 per student and raises the average student's mathematics achievement 10 points, the value of those 10 points can be estimated in terms of the students' increased incomes in adulthood. This is one way (although not the only way) of determining whether the school improvement is a good investment. If policymakers and researchers are interested in making such evaluations, it is crucial that the same test be given to both school-age children and adults or, alternatively, that when different tests are given they can be "scaled" to each other to enable direct comparisons between scores across different tests.

Which skills should be tested depends upon the specific interests of the policymakers, but in almost all cases they will include basic reading comprehension and mathematics. In addition, if the interest exists and resources are available, it would be useful to test for the following skills: science, health knowledge, innate intelligence, reading comprehension in another language, values and behavioral norms, agricultural knowledge, and basic practical knowledge. Testing intelligence is controversial, but there are several commonly used tests for this purpose, including the Raven's Progressive Matrices test. The skills covered

need not be limited to skills taught in the school curriculum; indeed, it may be interesting to see how people acquire skills that are not taught in school.

In general, the tests should be quite short; 10–20 questions should be sufficient for most purposes. In some cases it may be useful to have a basic test and an advanced test, allowing only people who achieve a “passing” score on the basic test to take the advanced test. The length of the test or tests should be thoroughly discussed with the consultant hired to develop the test. Many testing consultants are accustomed to long tests that are administered in classroom settings with few time constraints, so it may be necessary to impress upon the consultant the impracticality of administering long tests in a survey setting.

One yardstick for measuring school performance is how well students acquire the skills that the national curriculum is designed to teach. This implies that the tests should be designed to fit that curriculum. However, there may be reasons not to design tests this way. First, the curriculum itself may be outdated or irrelevant, and thus have little impact on many socioeconomic outcomes. Second, many skills of interest, such as agricultural skills or knowledge about health, may not be taught in schools. Third and finally, it is not clear that the precise skills that make workers more productive are the skills taught in schools. In general, the national curriculum may be a good place to start, but serious thought should be given to measuring other skills as well.

Perhaps the most difficult aspect of administering achievement tests as part of a household survey is finding a good place to administer the tests. If the tests are administered only to children currently in school and the data are collected when schools are in session, the tests can be administered in schools.<sup>23</sup> However, if adults or children not currently enrolled in school are to be tested, they must be tested either in the home or in some community center.

Testing in a community center has several advantages. There are minimal disturbances and it is easier than in other settings to prevent test takers from assisting each other. Test administrators have relatively more control over factors that may affect the test, such as adequate lighting and enough tables to write on. And there should be relatively few comparability problems with other tests that are administered in a similar setting.

The main disadvantage of testing in a community center is that it is difficult to get people—especially

adults—to come take the tests. One way to reduce this problem is to compensate subjects for being tested, but even this may not work with many adults. Although some testing specialists may have serious reservations about testing done in respondents’ homes, there are several measures that can be taken to minimize the problems involved. For further practical advice see UNNHSCP (1989) and Wagner (1993).

#### **ADDITIONS TO THE HOUSEHOLD QUESTIONNAIRE.**

Some additional questions to expand Part B of the standard module are given in the expanded questionnaire module page, to which the following comments apply.

**B11–B13.** These questions can be used to investigate how important textbooks are under different circumstances. These circumstances include whether the child owns a full set, whether the child can take the textbooks home (if they are provided by the school), whether the books are shared with others, and whether they are new or used. In principle, similar questions can be asked about other student-specific inputs such as exercise books, slates, pens, and pencils—but it is not clear whether it is feasible or worthwhile to collect detailed information on these items. Which items to include in detail may depend on the circumstances in a particular country.

**B14.** In principle, homework is one input into a learning production function, but it is clearly endogenous and it would be hard to find instrumental variables that can be excluded from the production function. If such instrumental variables can be found, it might be worthwhile to add a question about whether the child was assisted by an adult household member, as well as perhaps a question on the number of hours of tutoring the child received per week—distinguishing between assistance received from family members and assistance received from paid tutors.

**B16–B18.** These questions gather information on participation in school feeding programs in order to yield descriptive information on who benefits from these programs. More ambitiously, it might be possible to measure the impact of school feeding on educational outcomes such as attendance and test performance, or even on nutritional outcomes such as height and weight.

*School and Teacher Questionnaires*

As explained in the second section of this chapter, most methods for estimating the determinants of educational outcomes require information on the schools that are available for children to attend. LSMS surveys in the 1980s and 1990s have typically relied on information collected in the community questionnaire—information that has been much too brief. It is necessary to collect a large amount of information on school quality to reduce problems of omitted variable bias. Thus in the case of a standard, full-size multitopic household survey it is recommended that a school questionnaire be used for each local primary or secondary school. This may not be feasible in urban areas, so a reasonable rule of thumb is to take the five most commonly attended primary schools, the three most commonly attended lower secondary schools, and the three most commonly attended upper secondary schools. The basic school questionnaire, introduced in the third section of this chapter, is provided in Volume 3 of this book. Roughly speaking, it should only take about 30 minutes to complete for one school.

**SCHOOL QUESTIONNAIRE.** The following notes apply to the school questionnaire:

**A3.** In some countries schools may be combinations of these categories, such as a combined lower and upper secondary school. Additional codes for such cases should be created as needed.

**A6.** Information on when the school first opened is useful for analyzing the nonrandom placement of schools, as in Pitt, Rosenzweig, and Gibbons (1993).

**A8–A9.** If the number of shifts (Question A8) or the amount of class time per day (Question A9) varies by grade, extra columns should be created, as in Question A7, to get the information by grade.

**A9.** Class time per day is a critically important input into the learning production function. It also measures the opportunity cost of children's time spent in school.

**A10.** If many schools are open for fewer than the standard number of weeks per year, it may be useful to add a question on the reasons why this is the case.

**A11.** The answers to this question may vary by grade. If so, separate answers should be allowed for each grade. Also, if two or more languages are used for certain subjects, codes should be developed for combinations involving more than one language.

**B1–B5.** An alternative to these questions, which would involve more resources, is to administer teacher questionnaires, an example of which is discussed below. In such cases it would probably be necessary to retain only Questions B1 and B2 in the school questionnaire.

**B3.** In some countries certain teachers may have an explicit credential that others do not have. In such cases this question should ask specifically about that credential.

**C3.** A legible blackboard is any blackboard that can be used. It is possible to draw a finer distinction by asking whether a blackboard is sufficiently legible to be seen by all students or only by some students (either because some students have poor eyesight or because they sit at the back of the classroom).

**D2.** In countries in which several languages are spoken, information on the number of books in the school library should be collected separately for each language.

**D3.** In most countries, the question on science laboratories need not be asked in primary schools.

**E1–E4.** These questions apply only to school levels that have national examinations. In some cases the school principal may be able to provide more detailed information, such as how many students scored at several different levels.

**F1.** Obtaining accurate figures on school fees is in practice much more difficult than might be expected, since there are many different fees and the line between voluntary and mandatory is often unclear. It is best to list as many fees as possible by their exact names. If there is substantial variation by grade, it is best to obtain this information separately for each grade.

**F7–F9.** School uniforms can be a major cost to households, so relaxing the requirement to purchase a uni-

form may greatly reduce the real price of schooling for poor households.

**TEACHER QUESTIONNAIRE.** If the main focus of the survey is on education, it is important to collect information on individual teachers using a teacher questionnaire, a prototype of which is provided in Volume 3. In most countries teachers should be able to fill out the questionnaires themselves—vastly reducing the amount of time interviewers need to spend administering these questions. Yet even when teachers can fill these questionnaires out, the interviewer should check all of them on the spot to see if any errors have been made, and then ask the teachers to fix those errors.

If a school has more than, say, 10 teachers, it may be useful to choose two or three teachers per grade at random and administer the questionnaires only to them. One way to choose teachers at random is to ask all of them their dates of birth and then select the two or three with the earliest dates of birth within the calendar year.

A final possibility to consider when collecting information from individual teachers is to administer achievement tests to the teachers. In many developing countries, teachers' knowledge of the curriculum has serious deficiencies, and one would expect that the students of weaker teachers would do worse in school, other things being equal. However, this information can be quite difficult to collect. Teachers may resist being tested because they may think that the results will be used against them. Assurances that the data are being collected purely for research purposes may prove unconvincing (though not requiring the teachers within a given school to write their names on the tests might improve cooperation). Thus, if everything else is in place for the survey, it is worthwhile to explore the possibility of testing teachers, but if significant resistance is encountered, it may be best to drop this in order not to jeopardize the rest of the survey.

The following comments apply to the teacher questionnaire:

**Q12–16, Q17–21.** If achievement tests are being administered to students on subjects other than mathematics and reading, similar sets of five questions can be added for those subjects.

**Q34–35.** If these tasks are primarily done during class time, Questions 34 and 35 should be part of the previous set of questions (Questions 24–33).

**Q39–40.** If achievement tests are being administered to students on subjects other than mathematics and reading, it may be useful to add similar questions for those subjects.

**Q43–45.** If a new kind of in-service teacher training program is being used, the questions should specifically ask whether this training was of that type.

## Notes

The author would like to thank Jere Behrman, Kimberly Cartwright, Margaret Grosh, Eric Hanushek, Xiaoyan Liang, Marlaine Lockheed, Andrew Mason, Raylynn Oliver, Maris O'Rourke, Harry Patrinos, Julie Anderson Schaffner, and T. Paul Schultz for comments on earlier drafts.

1. Even tuition is not always paid. In some countries a child may attend school temporarily while the teacher or principal waits for the child's parents to pay. If the parents never pay, their children are usually forced to withdraw from the school (although sometimes poor children are allowed to stay).

2. There is one important exception to this statement. Distance to schools can directly affect learning if long distances lead to more absences or increased tardiness. This will be discussed in the next section of this chapter.

3. For material inputs that are child-specific, such as textbooks, an additional issue is whether the school should provide them or households should be required to purchase them. For most purposes this can be thought of as a price question, where provision by the school implies a price of zero. However, when schools provide these inputs, rationing may well take place, complicating efforts to understand the impact of such a policy.

4. Distance education can be viewed as a pedagogical technique because teachers turn on the radio, may participate as directed by the radio program, and are expected to follow up when the program is finished.

5. In this case, reduced school attainment may not be cause for alarm. It may be that high-paying government jobs were generating private rates of return to higher education that exceeded social rates of return.

6. Information about on-the-job training is best collected in the employment module of a multitopic household survey. See Chapter 9 for details.

7. Collecting information on schoolchildren's values may also be useful, but to the author's knowledge this has never been done for a representative sample of children in a developing country.

8. One estimation issue here is that the education variable may be endogenous. Another issue is measurement error in the educa-

tion variable. Both problems lead to biased estimates, and both are discussed further in the next section of this chapter.

9. A local school can be defined as any school within walking distance (more generally, within daily commuting distance) of a household.

10. Recall that causal factors are exogenous if they cannot be affected by household behavior (choices).

11. In theory, there could also be a production function for values, but to the author's knowledge no one has attempted to estimate such a relationship and interpret it as a production function.

12. There are two exceptions to this statement. First, distances to local schools can vary across households. Second, Hausman and Taylor (1981) provide a general method that, under certain conditions, can be used to estimate school effects, community effects, or both. However, to the author's knowledge, this method has never been used to analyze educational outcomes, and, in practice, the assumptions required for it to be valid may not be plausible.

13. Two implicit assumptions are made here: that the distance to the school attended is exogenous and that each child attends school nearly every day (which implies a fixed opportunity cost of travel time for each year of school enrollment). Both of these assumptions can be questioned.

14. This raises the issue of how to collect accurate data on distances. Households' estimates may not be very accurate, which will lead to bias due to measurement error. One recent development in survey design involves use of global positioning system technology, which can measure the latitude and longitude of any location within about 100 feet. If readings are taken for dwellings and local schools, precise distances (by air) can be calculated.

15. It would be possible to evaluate student loans using cross-sectional data in the context of a randomized evaluation. This general approach is discussed below.

16. This ignores grade repetition. If repetition is common, years of schooling should be used rather than the grade attained.

17. Note that assigning an interpretation to the parameter estimates associated with school management variables is risky, since the magnitude of their effect depends on the completeness of the data on material inputs, teacher characteristics, and pedagogical practices.

18. Alternatively, a reduced form relationship could be estimated by including information on local health clinics and medicine prices when estimating the determinants of educational outcomes. In this case, it is not even necessary to include data on child health, because it is substituted out of the reduced form relationship.

19. Equations for higher grades could be added in equation 7.1, but this has not been done in order to keep the exposition simple.

20. If previous test scores can be obtained from other sources, panel data are not needed. Unfortunately this is rare, and matching individuals from the two data sources can be very difficult. For

example, matching children from the 1990 Jamaican LSMS survey to national examination records proved impossible.

21. Consider two people—one who attended only primary school many years ago and another, currently in upper secondary school, who repeated one year of lower secondary school. Assume that neither has been an apprentice or had any technical education or training. Using the 1987–88 Ghana Living Standards Survey for comparison, the number of questions asked of the first person increases from 13 to 15, while the increase for the second person is from 17 to 26. The second person's increase is due mostly to four questions on grade repetition and two questions on the age and year when the person entered first grade. (This comparison ignores the questions in Part D, which are asked only once for each household and only if the household has children of school age.)

22. The main exception to this recommendation is that anthropometric data need not be collected in a country with very low rates of malnutrition. If the incidence of malnutrition is unclear, anthropometric data should be collected.

23. This does not mean that all children in a given school will be tested; if only a few sampled children attend a given school, they can be asked to remain at school after classes are over or permission can be obtained to withdraw them during the school day (as was done in the 1990 Jamaican LSMS survey).

## References

- Alderman, Harold, Jere Behrman, Victor Lavy, and Rehka Menon. 1997. "Child Nutrition, Child Health and School Enrollment: A Longitudinal Analysis." Policy Research Working Paper 1700. World Bank, Development Research Group, Washington, D.C.
- Becker, Gary. 1995. "Human Capital and Poverty Alleviation." HRO Working Paper 52. World Bank, Human Resources Development and Operations Policy Vice Presidency, Washington, D.C.
- Behrman, Jere, and Anil Deolalikar. 1991. "School Repetition, Dropouts and the Returns to School: The Case of Indonesia." *Oxford Bulletin of Economics and Statistics* 53 (4): 467–80.
- Burtless, Gary. 1995. "The Case for Randomized Field Trials in Economic and Policy Research." *Journal of Economic Perspectives* 9 (2): 63–84.
- Del Rosso, Joy, and Tonia Marek. 1996. *Class Action: Improving School Performance in the Developing World through Better Health and Nutrition*. Washington, D.C.: World Bank.
- Dow, William, Paul Gertler, Robert Schoeni, John Strauss, and Duncan Thomas. 1997. "Health Care Prices, Health, and Labor Outcomes: Experimental Evidence." RAND Labor and Population Working Paper 97-01. Santa Monica, Cal.

- Gertler, Paul, and Paul Glewwe. 1990. "The Willingness to Pay for Education in Developing Countries: Evidence from Rural Peru." *Journal of Public Economics* 42 (3): 251–75.
- . 1992. "The Willingness to Pay for Education for Daughters versus Sons: Evidence from Rural Peru." *World Bank Economic Review* 6 (1): 171–88.
- Glewwe, Paul. 1999. *The Economics of School Quality Investments in Developing Countries: An Empirical Study of Ghana*. London: Macmillan.
- Glewwe, Paul, and Hanan Jacoby. 1994. "Student Achievement and Schooling Choice in Low-income Countries: Evidence from Ghana." *Journal of Human Resources* 29 (3): 843–64.
- . 1995. "An Economic Analysis of Delayed Primary School Enrollment in a Low-income Country: The Role of Early Childhood Nutrition." *Review of Economics and Statistics* 77 (1): 156–69.
- Glewwe, Paul, Hanan Jacoby, and Elizabeth King. Forthcoming. "Early Childhood Nutrition and Academic Achievement: A Longitudinal Analysis." *Journal of Public Economics*.
- Glewwe, Paul, Michael Kremer, and Sylvie Moulin. 1999. "Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya." World Bank, Development Research Group, Washington, D.C.
- Glewwe, Paul, Margaret Grosh, Hanan Jacoby, and Marlaine Lockheed. 1995. "An Eclectic Approach to Estimating the Determinants of Achievement in Jamaican Primary Education." *World Bank Economic Review* 9 (2): 231–58.
- Grosh, Margaret, and Juan Munoz. 1996. *A Manual for Planning and Implementing the Living Standards Measurement Study Survey*. Living Standards Measurement Study Working Paper 126. Washington, D.C.: World Bank.
- Hanushek, Eric. 1986. "The Economics of Schooling." *Journal of Economic Literature* 24 (3): 1141–77.
- . 1992. "The Trade-off between Child Quantity and Quality." *Journal of Political Economy* 100 (1): 84–117.
- . 1995. "Interpreting Recent Research on Schooling in Developing Countries." *World Bank Research Observer* 10 (2): 227–46.
- Hausman, Jerry, and William Taylor. 1981. "Panel Data and Unobservable Individual Effects." *Econometrica* 49 (6): 1377–98.
- Heckman, James J., and Jeffrey A. Smith. 1995. "Assessing the Case for Social Experiments." *Journal of Economic Perspectives* 9 (2): 85–110.
- Ilon, Lynn. 1992. "School Unit Cost Study: Jamaica." State University of New York, Graduate School of Education, Buffalo, N.Y.
- Jacoby, Hanan. 1993. "Shadow Wages and Peasant Family Labor Supply: An Econometric Application to the Peruvian Sierra." *Review of Economic Studies* 60 (4): 903–21.
- . 1994. "Borrowing Constraints and Progress through School: Evidence from Peru." *Review of Economics and Statistics* 76 (10): 151–60.
- Knight, John, and Richard Sabot. 1992. *Educational Performance of the Poor: Lessons from Rural Northeast Brazil*. New York: Oxford University Press.
- Kremer, Michael, Sylvie Moulin, David Myatt, and Robert Namunyu. 1997. "Textbooks, Class Size and Test Scores: Evidence from a Prospective Evaluation in Kenya." Massachusetts Institute of Technology, Department of Economics, Cambridge, Mass.
- Lockheed, Marlaine, and Adriaan Verspoor. 1991. *Improving Primary Education in Developing Countries*. New York: Oxford University Press.
- Murnane, Richard, John Willett, and Frank Levy. 1995. "The Growing Importance of Cognitive Skills in Wage Determination." *Review of Economics and Statistics* 77 (2): 251–66.
- Newman, John, and Paul Gertler. 1994. "Family Productivity, Labor Supply, and Welfare in a Low-Income Country." *Journal of Human Resources* 29 (4): 989–1026.
- Newman, John, Laura Rawlings, and Paul Gertler. 1994. "Using Randomized Control Designs in Evaluating Social Sector Programs in Developing Countries." *World Bank Research Observer* 9 (2): 181–201.
- Pitt, Mark, Mark Rosenzweig, and Donna Gibbons. 1993. "The Determinants and Consequences of the Placement of Government Programs in Indonesia." *World Bank Economic Review* 7 (3): 319–48.
- Pollitt, Ernesto. 1990. *Malnutrition and Infection in the Classroom*. Paris: United Nations Educational, Scientific, and Cultural Organization.
- Selden, Thomas, and Michael Wasylenko. 1995. "Measuring the Distributional Effects of Public Education in Peru." In D. van de Walle and K. Nead, eds., *Public Spending and the Poor: Theory and Evidence*. Baltimore, M.D.: Johns Hopkins University Press.
- Strauss, John, and Duncan Thomas. 1995. "Human Resources." In J. Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics: Volume 3*. Amsterdam: North Holland.
- UNDP (United Nations Development Programme). 1990. *Human Development Report*. New York: Oxford University Press.
- UNNHSCP (United Nations National Household Survey Capability Program). 1989. *Measuring Literacy through Household Surveys*. New York: United Nations.
- Wagner, Daniel. 1993. *Literacy, Culture, and Development: Becoming Literate in Morocco*. New York: Cambridge University Press.
- West, Edwin. 1996. "Education Vouchers in Practice and Principle: A World Survey." Human Capital Development Working Paper 64. World Bank, Washington, D.C.

World Bank. 1990. *World Development Report—Poverty*. New York:  
Oxford University Press.

———. 1995. *Priorities and Strategies for Education*. Washington, D.C.  
———. 1997. *Primary Education in India*. Washington, D.C.



# 8

## Health

*Paul J. Gertler, Elaina Rose, and Paul Glewwe*

Health is a critical factor in the development of any country, for two reasons. First, health status is a key indicator of a population's welfare (Sen 1985). Second, improving the health status of the population leads to greater economic productivity (Strauss and Thomas 1995). In almost every country, governments play an active role in the health sector, and there are sound economic reasons for their doing so. The main reason is that there are often "market failures" in the purchase and provision of health services, including incomplete health insurance markets and imperfect information on the part of the population about the nature of health problems and their treatments.

To reduce the adverse effects of these market failures, policymakers need information on health outcomes, on household behavior that affects those health outcomes, and on the effects of government policies on both outcomes and behavior. In recognition of this policy need, almost all past LSMS surveys have included a health module, and the data from these modules have been used to study several important health sector policy issues (for example, the study of willingness to pay for health care by Gertler and van der Gaag 1990). Although data from the LSMS health module have been put to productive use, it is time to re-examine the module to see how well the information it collects meets the current challenges in the health sector.

The purpose of this chapter is to help survey designers develop a health module for LSMS and similar household surveys that will yield the information needed by health sector policymakers. The first section discusses key health policy questions. The second section outlines the data needed to analyze those policy questions. The third section introduces three modules designed to collect these data. (The modules are pro-

vided in Volume 3 of this book.) The fourth section presents annotations to these modules.

### **Major Health Policy Issues in Developing Countries**

Every government intervenes in the health sector to some extent, but the nature and size of those interventions varies from country to country. Economic theory gives three main reasons for governments to intervene in the health sector. First, health is, in many respects, a public good. This means that an individual's good health benefits not only him or her but also other members of society. For example, when a person with a contagious disease decides to seek treatment, this reduces the probability that others will get that disease. Second, there are many market failures that prevent the effective use of available resources in the health sector. For example, when health insurance markets are incomplete, this means that many families whose members develop a serious illness either cannot afford to pay for treatment or, if they can afford to pay,

will incur major financial losses by doing so. Third, equity in health outcomes is an important social goal of many governments, and one way for governments to increase equity is to support health care services. For further discussion of the economic rationale for government intervention in the health sector see de Ferranti (1985), Hammer (1997), and Gertler and Hammer (1997).

These general reasons for government involvement in the health sector raise a large number of more specific questions concerning how governments should be involved. The first step in addressing these questions is to identify specific health policy issues. This section begins with a brief description of how health care is provided in developing countries and then reviews the most important policy issues concerning health in developing countries.

#### *Health Care Provision in Developing Countries*

Governments in developing countries provide health care services in many different ways. One common way in which governments are involved in the health sector is through public health programs. These include immunization programs, other programs to prevent the spread of highly infectious diseases, treatment programs for individuals already infected, and public education programs on such subjects as smoking, risky sexual activity, hygiene, nutrition, and preventative health care. The distinguishing characteristic of these programs is that they include activities designed to reach out to people in their communities, homes, and workplaces, as opposed to reaching only those people who visit health care facilities. Policymakers can use information collected in household surveys to make these programs more effective. For example, household survey data can reveal which households participate in or are otherwise affected by public health programs—and thus indicate whether those programs are reaching their intended beneficiaries.

Governments also provide preventative and curative medical care through publicly operated hospitals, clinics, and other health care facilities. The prices for these health care services are often heavily subsidized using general tax revenues. These subsidies are provided to reduce individuals' financial risk, to increase the access that disadvantaged groups such as women, children, the poor, and the elderly have to health care, and to improve the overall health status of the population. In financing and operating health facilities, govern-

ment officials have to make many decisions about the allocation and mobilization of resources, including the location of different types of facilities (such as hospitals or clinics), which services each type of facility should provide (for example, inpatient care, x-rays, surgery, drugs, family planning, and preventative care), the quality of these services, and the fees to charge for these services. Household surveys can collect much of the data needed to study the consequences of these different kinds of policy decisions. Such data are critical for deciding how to finance and operate public health care facilities. For example, policymakers can use household survey data to assess how health outcomes and utilization patterns would change if user fees were increased or if the quality of health care services were improved at publicly operated facilities.

Most countries' governments allow—and in some cases actively encourage—the private sector to deliver some health care services. In such countries, for the government to develop effective public sector policies, it needs accurate information on the extent to which the private sector complements or substitutes for the public sector in meeting government objectives. For example, it is useful to know the extent to which individuals switch from using the public sector to using the private sector in response to increases in user fees in the public sector, or from the private to the public sector when the quality of public sector care improves. Information gathered in household surveys can be used to analyze these relationships.

Finally, many governments are introducing mandatory social health insurance in the formal wage sector. This social insurance is typically financed through a payroll tax that goes into a fund used to pay for workers' medical care when they are ill. This fund insures workers against the financial cost of their illnesses or injuries and also reduces the government's health budget by requiring workers to contribute to the social insurance fund. However, when people are fully insured, they tend to use more health care than is socially optimal. This phenomenon is known as moral hazard. Knowing the likely extent of moral hazard helps policymakers to set copayments and premiums at a level that ensures the financial viability of the insurance fund. Moral hazard is measured by the price elasticity of demand, which can be estimated using household survey data.

These different ways in which governments intervene in the health sector do not necessarily achieve

their intended objectives, and they may even have negative consequences. Indeed, in some developing countries, the overall health care system functions poorly and thus does little to improve the overall health status of the population (World Bank 1993; Peabody and others 1997). The ability of the government to improve its policies depends on the accuracy and timeliness of the information that the government has at its disposal regarding current circumstances in the health sector and the likely impact of different interventions on households' choices and outcomes. The rest of this section reviews in more detail the most important policy issues in the health sector.

### *Current Health Policy Concerns*

Health policy concerns may be grouped by category of disease or by category of people with health problems (for example, children, women, or the elderly). Yet for the purpose of trying to understand the impact of different policies, it is more useful to divide policy issues according to the type of policy. This section divides health policy concerns into eight different categories:

- Assessment of health problems and associated behavior.
- Equity in health status and in access to health services.
- Provision of public health programs and services.
- Pricing policies for health services.
- Maintenance and improvement of the quality of health services.
- Regulation of privately provided health services.
- Health insurance policies.
- The impact of health on other socioeconomic outcomes.

These categories are discussed in detail in the following paragraphs.

**ASSESSING HEALTH PROBLEMS AND MEASURING ASSOCIATED BEHAVIOR.** The starting point for discussing almost any health policy issue is a good understanding of current health problems and the types of behavior associated with those problems. Thus accurate assessment of the current situation can be thought of as a distinct health policy concern. For the population as a whole and for particular disadvantaged groups—the poor, women, children—policymakers need baseline information on the level and distribution of health problems to identify which groups are most at risk and to provide benchmarks for judging

the effectiveness of government interventions. In particular, information is needed on the level, distribution, and causes of both child and adult mortality within the population, on the incidence of specific serious diseases among different demographic and socioeconomic groups, and on the extent to which people are unable to carry out their usual activities because of poor health.

In some countries access to medical care may not be the crucial issue. The key to the population's health status may be household-level factors such as water supply, sanitation, waste disposal and cooking practices, or individual-level behavior such as diet, infant feeding practices, exercise, use of seatbelts, tobacco use, alcohol consumption, and sexual behavior. The relative importance of each of these factors and kinds of behavior varies from country to country. Policymakers in each country need to know the prevalence of different kinds of health-related behavior, the extent to which the population is aware of the impact that these kinds of behavior can have on health, and which groups within the population are engaging in behavior that can adversely affect their health.

**EQUITY ISSUES.** Many policymakers and governments are concerned not only with the overall incidence of health problems and the provision of health services but also with the distribution of both health problems and the use of health services among the population. In almost every country the poor are more likely to suffer from bad health and less likely to receive health care services. While governments cannot fully compensate poor households for their lack of income, they can implement or adjust health care policies in ways that will help the poor increase their use of such health care services as immunization, prenatal care, and medical treatment for specific illnesses.

The first thing that policymakers need is information on the incidence of specific health problems broken down by income groups in the population. This will give them an idea of the degree of inequality that exists under current policies. The second type of information that policymakers need is who receives medical care under the current system, disaggregated by the type of care provided (for example, preventative, curative, or prenatal), the type of provider (for example, public, private, or traditional), and the level of care received (primary, secondary, or tertiary).<sup>1</sup> This information allows policymakers to examine the relation-

ship between differences in the use of health services among subgroups of the population and differences in health outcomes.

Equity involves more than just health status and who gets health care services. It also involves how much individuals and households pay for those services. Thus a third useful piece of information for analyzing equity issues is data on individual and household expenditures on health care. Households make private out-of-pocket expenditures on outpatient visits for preventative purposes (including prenatal care, immunizations, and checkups), on services during childbirth, on outpatient visits for curative purposes, on inpatient stays, on medicine, on travel to and from health care facilities, and on food during their inpatient stays. Policymakers would also like to know whether households increase their expenditures on private health care services in response to a reduction in publicly provided health care services.

A final equity issue is how government subsidies for health care are distributed across different socioeconomic groups. When the fee charged for a given health care service is lower than the cost of providing that service, the government is subsidizing that service. Policymakers need to know who really benefits from these subsidies, and this means they need to have an accurate picture of how public subsidies are currently being allocated. Benefit incidence analysis can be used to calculate the distribution of public subsidies among beneficiaries, giving governments the opportunity to assess whether the current distribution of subsidies is consistent with its policy priorities.

**PUBLIC HEALTH PROGRAMS.** One action governments can take to improve the health status of the population is to provide public services that directly improve people's health or alter people's behavior in a way that results in an improvement in their health. The best-known example of these services is child immunization campaigns, in which health workers visit homes and schools to immunize children against a variety of serious childhood diseases (such as diphtheria, pertussis, tetanus, measles, and polio). A second example is the opening of a new water treatment facility, which directly improves the health of the local people by increasing the quality of their public water supply. A third example is the provision of public garbage collection services, which results in more sanitary garbage disposal. Finally, there is the example of government

policies to reduce air pollution, which should lead to improved health, although there has been little research in this area. For each of these examples, policymakers need to know how changing the availability of the relevant government programs and services affects individuals' health and health-related behavior.

Another way in which the government can improve the health status of the general population is through public education campaigns that provide the public with information on the consequences of individual and household health-related behavior. Publicity campaigns can inform individuals about the hazards of activities such as smoking, alcohol consumption, and unsafe sex practices, as well as the benefits of activities such as breastfeeding, seat belt use, exercise, and good nutrition. Because there may be a long time-lag between the introduction of a campaign and any changes in the population's health, it is difficult to find direct evidence that public education campaigns lead to improvements in the population's health. It may be more useful to look at the effects of these campaigns on behaviors—for example, in the case of an anti-smoking campaign, by examining whether there has been a reduction in the number of smokers—or on people's knowledge of the risks associated with different types of behavior.

**PRICING POLICIES FOR HEALTH CARE SERVICES.** Many governments charge fees for publicly provided health care services to help finance the provision of these services. However, as the fee charged for any service increases, the demand for that service tends to decrease. There may also be switching from one kind of service to another; for example, an increase in the price of publicly provided services may lead people to switch from public services to private services. Changes in the use of health care services can have serious health implications, so policymakers need to know the net effect of a change in user fees on households' use of health care services. They would also like to know the effect of changing user fees on government revenues. This depends on the impact of price changes on the utilization of health care services—that is, the price elasticity of demand for these services.<sup>2</sup> More generally, policymakers would like to know the impact of changing the user fee charged for a particular health care service on the utilization of that and other health care services, on households' health care expenditures, on which service providers they use, and

on government revenues. Ideally, policymakers would like to have this information for each type of health care service provided by the government.

Another factor that may explain why certain groups have low utilization rates is that they may have to travel farther than other groups to reach the nearest health care facility. Distance can be thought of an additional “price” that households pay to use health facilities. The location of facilities can affect households’ choices of which provider to use, the extent to which households use any health care services, and the costs that households must incur to obtain treatment. Thus policymakers need to know how the location of public health facilities affects the utilization of those facilities, household expenditures on health care, the service providers that households choose, and government revenues from health care fees.

A final way in which governments can use pricing policies to influence health is by affecting health-related behavior. For example, governments can impose taxes on the purchase of goods such as cigarettes and alcohol, which is likely to reduce the consumption of these goods. To ensure that such a policy is having the desired effect, a government needs to know how the imposition of (or change in) such taxes affects the consumption of these goods as well as how it affects household expenditure patterns and government revenues.

**QUALITY OF HEALTH CARE SERVICES.** Another determinant of the utilization of health care services is the quality of the care provided. If certain groups within the population rarely use health care services, it may be because the quality of these services is low. In general, households will increase their use of services if the quality of the services increases while the price remains unchanged. This implies that household expenditures on health care services (and, consequently, government revenues from user fees) will increase in response to an improvement in the quality of health care facilities. Thus the quality of the care provided by a facility has both a direct effect on health and an indirect effect through changing the utilization of services. Policymakers need to know how changing the quality of care affects utilization patterns, households’ choice of provider, household expenditures on health care, government revenues from user fees, and, ultimately, health status.

A particularly important issue regarding the quality of health care services is whether prices and quali-

ty can be simultaneously increased in such a way that the increased revenues from the higher prices will pay for the increase in the quality of the services. In some cases the increase in utilization rates due to improved quality may outweigh the reduction in utilization rates due to the increased prices. If this is the case, utilization will increase, or at least will not decrease—thus improving the health status of the population. Whether this fortuitous result can actually be realized has been examined in several recent studies (such as Litvak and Bodart 1993 and Peabody, Gertler, and Leibowitz 1998). Further research is needed to see how likely it is, and under what conditions, that price increases can finance improvements in quality without reducing health care utilization.

**POLICIES REGARDING PRIVATELY PROVIDED HEALTH CARE SERVICES.** There are several reasons why private providers should be allowed, and perhaps even encouraged, to provide some types of health services. The main reason is that the private sector can often provide goods and services more efficiently than the government can. Yet allowing the private sector to provide certain types of health care services does not necessarily imply that the government has no role to play. For example, the government can play an important role in informing the general public about the efficacy of particular treatments for specific diseases and about the reputations of private providers.

Perhaps the most important role the government can play is to regulate private providers so the public is assured that private services provide a minimum level of competence and quality. This can be done in several ways. First, governments can implement a system to train and license health care providers. Second, governments can monitor the quality of health care facilities on a regular basis, focusing on both the quality of the facilities (such as whether they have the proper equipment and whether the environment in which they operate is hygienic) and the process of care (such as whether proper prenatal care practices are being followed). A third area in which the government can be involved is in the regulation of pharmaceuticals; governments must decide which drugs should be banned, which should be available only by prescription (and who should be allowed to prescribe them), and which can be sold over the counter. Fourth, the government can regulate the prices charged by private providers, although some economists advise against

regulation of prices in the private sector. A fifth way in which the government can usefully intervene is in the market for private insurance, where it may want to regulate certain practices that can be detrimental to consumers, such as the refusal to provide coverage to particular individuals or groups.

Another reason why governments need information on private health care facilities is to know the extent to which individuals change from the public to the private sector in response to policy changes at public health facilities. A related issue is the extent to which private providers adjust the price and quality of their services in response to changes in the price and quality of public services. Performing such analysis requires that information be collected from both private and public providers using a health facility questionnaire.

#### **HEALTH INSURANCE AND EMPLOYER-PROVIDED**

**HEALTH BENEFITS.** Health insurance is a program or contract in which part or all of the cost to an individual of obtaining medical treatment is paid for by the insurer. This insurance may be provided by the person's employer or by another source such as the government or private providers. Health benefits are provided by employers to their employees and thus are not available to self-employed workers. Such benefits consist of sick pay, sick leave, and maternity pay and leave. Some employers provide health care at the place of employment, but this is best understood as a type of health insurance rather than a health benefit.

Insurance coverage and employer-provided health benefits are less common in developing countries than in industrial countries, but they do exist in many low- and middle-income countries. For example, 14 percent of Indonesians working in the wage sector are covered by health insurance (World Bank 1993). Also, 9 percent of all Jamaicans and about 25 percent of all Brazilians have health insurance (Gertler and Sturm 1997; Lewis and Medici 1995). In developing countries as a whole, formal health insurance is growing and will become increasingly important as these countries complete the demographic transition whereby chronic diseases—which are expensive to treat—become a bigger problem than infectious diseases—which can be prevented and treated at low cost. Similarly, as the economies of low- and middle-income countries develop, there is likely to be an increase in the number of both public and private employers that offer health benefits to their workers.

Thus there is an increasing need for governments in developing countries to think about the roles played by insurance and employer-provided health benefits in determining health outcomes.

The increase in insurance coverage will bring about several changes in the health sector in developing countries. First, this increase is likely to raise utilization rates because individuals covered by health insurance pay less for treatment than do individuals who are not covered. While this outcome may seem unambiguously desirable, it is possible that the moral hazard problem associated with insurance will lead to overuse of health care services. Second, insurance coverage will affect household expenditures on health care by increasing the demand for health care. Third, individuals with health insurance may reduce their rate of (precautionary) savings because they no longer have to save to be able to pay for large, unexpected medical costs.

Given the increased availability of insurance, and its likely impact on health behavior and outcomes, policymakers need to know which individuals and households are covered by health insurance, whether through public, private, or employer-provided plans. Detailed data can be used by researchers to analyze the impact of health insurance schemes on the use of health services. In particular, researchers need information on the benefits provided by each insurance plan, the services covered, and the copayments, deductibles, and benefit caps associated with each plan. This will enable them to study how health insurance coverage affects health care utilization, household spending on health care, government revenues, and other household behavior such as consumption expenditures and savings. Over time, the increased availability of insurance may mean that government subsidies for health care can be reduced. However, these subsidies should not be eliminated altogether because there are other reasons to subsidize health care, such as the need to subsidize public goods and the desire to promote more equitable health outcomes.

#### **HEALTH STATUS AND OTHER SOCIOECONOMIC**

**OUTCOMES.** As mentioned in the introduction, an individual's health status can affect his or her economic productivity (Deolalikar 1988; Strauss 1986; Strauss and Thomas 1995; Dow and others 1997). It can also affect education outcomes (Behrman and Lavy 1992; Glewwe and Jacoby 1995; Glewwe, Jacoby, and King Forthcoming) and consumption and savings decisions

(Kochar 1995, 1997; Gertler and Gruber 1997). When governments decide how many resources to devote to health, they should account not only for the direct benefit of increasing the population's health status but also for the impact of health on other socioeconomic outcomes. Thus governments need to know how the health status of the population affects worker productivity, education outcomes, and consumption and savings. LSMS surveys are well suited for addressing these issues because they collect data on all of these topics, as seen in the other chapters of this book. Moreover, much of the data collected in LSMS surveys, such as information on worker productivity and education outcomes, is collected at the individual level (as opposed to the household level), which allows for a more disaggregated analysis of the relationship between health and other socioeconomic outcomes.

### **Data Needs for Policy Analysis**

This section describes the data that policymakers and researchers need to address the health policy issues presented in the previous section. The first subsection discusses how to assess health status using data from a household questionnaire. A thorough discussion is necessary because health status is complex and difficult to measure. Data on health status can be used not only for assessing current health problems but also for analyzing the impact of health status on other socioeconomic outcomes. The second subsection discusses how to use a household questionnaire to collect data on health-related behavior, the utilization of health facilities, health expenditures, insurance status, and access to services. Household data on health status, the utilization of health facilities, and health expenditures are essential for examining equity issues.

The third subsection turns to the community, price, and facility questionnaires, examining what data are needed to analyze issues concerning pricing policies, quality of health services, public health programs, and regulation of private health care services. The fourth subsection provides a short discussion of how to use the data collected for policy analysis. The fifth subsection briefly examines two important sampling issues. The sixth and final subsection links the policy issues discussed in the first section of this chapter to the short, standard, and expanded versions of the draft health module presented in Volume 3. Specifically, the sixth subsection presents 37 policy

questions (drawn from the discussion in the first subsection) and indicates which parts of the health module, as well as which other sources of relevant data, yield the information needed to answer each of these questions.

This section, and the rest of this chapter, assumes that the survey designers have already decided to collect health data as part of an LSMS-type multitopic household survey. Another alternative might be to design a survey devoted exclusively to health issues. Box 8.1 discusses the advantages and disadvantages of each option.

### *Assessing Health Status Using the Household Questionnaire*

Health status can be very difficult to measure in the context of a household survey (McDowell and Newell 1996; Stewart and Ware 1992). The ideal survey would employ doctors, nurses, or other health professionals to give each household member a complete health examination, including laboratory tests. However, the expense and logistical complications of doing this would be very high. Thus only household surveys that focus almost exclusively on health issues, as opposed to multitopic surveys such as LSMS surveys, can devote the resources required to collect such complete health data. Because of the costs and complications involved, most developing countries collect very little data on the health status of the general population; what these countries do collect tends to be limited to mortality data (which may or may not include information on cause of death) and, in some cases, anthropometric data. In most cases multitopic surveys must organize a new data collection exercise to gather the data needed to analyze the health policy questions discussed in the first section of this chapter.

There are two issues to consider when measuring health status. First, health is multidimensional. Nutritional status, morbidity, physical functioning, and mental functioning (mental health and cognitive ability) reflect different aspects of a person's health. These different dimensions of health may respond differently to policy changes and may have different effects on other important outcomes such as an individual's earnings and productivity and even his or her sense of well-being. Second, because people's perceptions of their own health are likely to be related to their education, occupation, and household income, self-reported information on health obtained from house-

**Box 8.1 Collecting Health Data in a Multitopic Survey or in a Single-Topic Survey**

The first decision that survey designers need to make is whether to collect health data as part of a multitopic household survey such as an LSMS survey or as part of a survey devoted almost exclusively to health issues, such as a Demographic and Health Survey. (For information on these surveys see the Demographic and Health Survey website at <http://www.macoint.com/dhs>.) The two main advantages of specialized health surveys are the detailed information that they can collect on health status and their large sample sizes, which will yield data that can be used to study a wide range of health problems, including rare events such as maternal mortality. However, health surveys have the serious disadvantage that they rarely collect the basic socioeconomic information needed to describe how health varies according to household income. This is primarily due to two constraints. First, collecting household income information adds to the cost of the survey. Second, collecting such information tends to put an unreasonable burden on households' time. Facility-based health surveys have the additional disadvantage that they cannot obtain information on the health status of the general population because they can collect information only on individuals who visit health facilities, and not on individuals who never visit any health facility.

Multitopic household surveys do not suffer from these disadvantages because they collect data not only on health but also on many other topics, such as employment and earn-

ings, education, and income and expenditures. Yet multitopic surveys have two serious disadvantages. First, the information that they collect on health status is limited, and much of what they do collect is reported by the respondents themselves rather than recorded by a trained observer. This reliance on self-reporting by respondents is cause for concern because such data tend to be less reliable. Second, for reasons explained in Chapter 1, multitopic surveys such as LSMS surveys tend to have samples that are usually no more than 5,000 households, which is too small to yield data for calculating most disease-specific measures of health, such as levels of coronary heart disease and cancer. (An important exception to this point is diarrheal disease in children; this is well defined and highly prevalent in most developing countries, so LSMS-type surveys should collect data on it.) On the other hand, a small multitopic household survey is a good vehicle for collecting summary measures of health status, such as anthropometric measurements of children, body mass index of adults, and physical and cognitive abilities of the elderly. These summary measures provide good benchmarks for measuring changes in overall health status over time. They are also useful for analyzing how the financing and accessibility of health care services affect health outcomes. Finally, they can be used as summary measures of health status in analyses of the impact of health on labor market productivity, educational performance, and other socioeconomic outcomes.

hold surveys must be interpreted with caution. In most previous LSMS surveys, each individual's health status was assessed by asking whether he or she was sick or injured at any time during the previous four weeks and, if so, whether his or her usual activities were limited by this illness or injury. Some LSMS surveys also asked respondents to self-report symptoms and to diagnose specific illnesses for themselves and their children. In contrast, and as explained more fully below, a better approach is to collect several objective measures of health status to avoid relying entirely on subjective (self-reported) data.

Self-reported measures of health status include general health status, limitations in daily activities, current morbidity, activities of daily living, and mental and emotional health. These will now be examined in turn. This will be followed by a brief discussion of several objective measures of health, including anthropometric status, mortality, directly observed physical functioning, clinical diagnosis of illness, information from medical tests, measures of cognitive functioning, and observed activities of daily living.

**SELF-REPORTED GENERAL HEALTH STATUS.** Self-reported general health status is an index of overall health based on the respondent's answer to the question, "In general, how is your health at this time?" The possible answers to this question are excellent, very good, good, fair, poor, and very poor. This measure of health status is correlated with future mortality, even after controlling for many other variables (see Idler and Benyamini 1997 and Ferraro and Farmer 1999). Even so, there are potential biases because respondents' answers tend to depend on their subjective standards of what constitutes "healthy" and on the extent to which they have had contact with the health system. For example, Dow and others (1997) reported that serious bias was revealed when this measure was used in models of labor supply.

**SELF-REPORTED LIMITATIONS IN DAILY ACTIVITIES.** Questions on limitations in daily activities include whether the respondent was able to perform his or her usual activities, the number of days during which his or her normal activities were limited, and whether he

or she was confined to bed because of illness. The main problem with these kinds of questions is that the responses to them will depend not only on the person's health but also on what his or her normal activities are and on how easily he or she is able to curtail those activities. The direction of the bias is uncertain. For example, if individuals who earn low wages tend to have more physically strenuous jobs, they are more likely to be constrained by a given illness than higher wage earners. On the other hand, since individuals who earn low wages may be less able to "afford" to rest when they are ill relative to those who earn higher wages, they may be less likely to limit their activities because of illness.

**SELF-REPORTED MORBIDITY.** Many household surveys, including several LSMS surveys, have asked respondents whether they are currently ill. For respondents who report being ill, some of the surveys have also asked what illness they were suffering from. Unfortunately, such self-reported information on morbidity may be biased by variation in respondents' perceptions and by differences in their knowledge of specific illnesses. It has often been found that the incidence of reported adult sickness increases with income and education. (Two examples using LSMS data are Schultz and Tansel 1997 and Dow 1996.) Because of this problem, self-reported data on symptoms and diagnoses for adults are of limited usefulness in assessing health status.

A similar problem arises with the way in which mothers report their children's illnesses. Using data from a health survey done in Peru, Sindelar and Thomas (1991) found that the probability of a mother reporting that her child has a respiratory illness increases with her education. On the other hand, the authors also found that the reporting of children's diarrhea declines with the mother's education, which suggests that reported diarrheal disease is more accurate than reported respiratory illness. Thus data from mothers on the incidence of diarrhea among their children may be reliable, but the value of asking mothers about their children's other illnesses or symptoms is doubtful.

**SELF-REPORTED ACTIVITIES OF DAILY LIVING.** Activities of daily living are derived from a series of questions regarding respondents' physical ability to carry out a number of activities. They are divided into

two categories. Intermediate activities of daily living consist of the following abilities: carrying a heavy load for 20 meters; sweeping the floor or yard; walking for 5 kilometers; drawing water from a well; and bending, kneeling, or stooping. Basic activities of daily living are the ability to bathe oneself, feed oneself, put on clothes unaided, stand up from a sitting position in a chair, go to the toilet unaided, and rise from sitting on the floor.

Activities of daily living are less subjective than other self-reported measures of health because they are well-defined, are not expressed in terms of respondents' normal activities, and do not require respondents to provide general opinions about their own health. Initially developed to study levels of disability among the elderly, these measures are used increasingly to study the health status of all adults. These measures of physical functioning have been tested extensively for reliability (consistency across tests and among different interviewers) and validity (consistency among individual assessments of different skills). In the United States, Jamaica, and Southeast Asia, they have been found to be reliable and valid self-assessments with a high degree of internal consistency (Andrews and others 1986; Guralnik and others 1989; Ju and Jones 1989; Strauss and others 1993; Ware, Davies-Avery, and Brook 1980). Moreover, they are the key measures of health status in the new U.S. Health and Retirement Survey (Wallace and Herzog 1995).

Activities of daily living have been used as dependent variables in many analyses of adult health and as explanatory variables in analyses of the consequences of ill-health. They are often used in studies of labor supply in the United States (for example, Bound and others 1991; Bound, Schoenbaum, and Waidman 1995; and Stern 1989) and have recently been used in Indonesia (Dow and others 1997). Gertler and Gruber (1997) used changes in activities of daily living to investigate whether families are able to insure their consumption against major illnesses. If activities of daily living are gathered using carefully worded questions, they will not be subject to the same reporting biases that are common in data on self-reported illness. They also have the advantage of being relatively easy to collect. Thus self-reported activities of daily living are an important measure of adult health status. In fact, a number of studies in both developed and developing countries have shown that, unlike with self-reported

morbidity, well-educated individuals in high brackets report fewer problems with activities of daily living than do poorer, less educated individuals (Strauss and others 1993; Gertler and Zeitlin 1996; Kington and Smith 1997).

**SELF-REPORTED MEASURES OF MENTAL OR EMOTIONAL HEALTH.** Self-reported measures of mental or emotional health can also be obtained from household surveys by asking respondents about how often they have experienced insomnia, fatigue, moodiness, impulsive anger, malaise, or sadness. While these measures have been validated in the United States and other developed countries, little evaluation has been done of their appropriateness for developing countries. More experience is needed on the feasibility and usefulness of collecting such data in developing countries.

**ANTHROPOMETRIC MEASURES.** Turn now to objective measures of health status. For adults, body mass index can be computed from data on weight and height, which are relatively easy to collect. Among adults, body mass index is associated with malnutrition and poor health. For children, nutritional status can be measured by standardized weight-for-height and height-for-age. This implies that weight and height data should be obtained for all household members in a multitopic household survey. Arm circumference and birth weight are additional indicators of health status. Chapter 10 discusses anthropometric measures in great detail, so there is no need for further discussion in this chapter.

**ADULT AND CHILD MORTALITY.** Although mortality is invariably a concern of policymakers, data on adult mortality is difficult to collect. In general, at least two surveys are needed to accurately measure adult mortality. The problem with relying on a single survey, which would require retrospective questions on adult household members who died in recent years, is that many adults live alone; deaths of adults who lived alone cannot be obtained retrospectively because the households in which they lived no longer exist. Also, adult mortality is a relatively infrequent event in households, so the samples needed to measure this precisely are much larger than those typically used in LSMS surveys. One exception is the use of sisterhood methods, which involve asking women whether any of their sisters have died in recent years. These methods

can be used to detect maternal deaths associated with childbirth (see Graham, Brass, and Snow 1989).

In contrast, data on child mortality data can be collected by obtaining fertility histories from mothers, which has been done in many previous LSMS surveys (see Chapter 15 for several examples). Because infant and child mortality rates are much higher than adult mortality rates, the sample sizes needed are not as large. Also, children rarely live alone, so panel data are not necessary. The use of these data to study infant and child mortality is discussed in detail in Chapter 15, and thus will not be considered further in this chapter.

**DIRECTLY OBSERVED ACTIVITIES OF DAILY LIVING.** One way to improve upon self-reported activities of daily living is to have a survey interviewer watch household members perform several relatively simple activities of daily living, such as standing in various positions, sitting up from a chair, and walking a short distance. Direct observation of activities of daily living often yields more accurate data than recording respondents' reports of activities of daily living. Directly observed activity of daily living measures have been used extensively in national health surveys in the United States and were recently used in the Matlab Health and Socioeconomic Survey in Bangladesh (Rand Corporation 1998).

**CLINICAL DIAGNOSIS.** Clinical diagnosis can be used to measure health status in a way that avoids many of the problems associated with self-reported illness. However, measurements obtained from clinical diagnoses involve substantial data collection problems. A sample of people who visit a health facility in a given day will be too small to be useful, which means that it is necessary to rely on health providers' memories of events for additional information.

If the survey budget can bear the substantial costs involved, some relatively simple medical tests can be conducted. For example, individuals' blood pressure and temperature can be measured by anyone who has been given a small amount of training, while lung capacity, which reflects people's long-term health status, can be evaluated relatively easily by nonmedical personnel who have been trained to use peak flow meters. Finally, finger-prick blood tests can detect anemia and micronutrient deficiencies, and HIV can be detected from saliva samples; both of these tests can be done by nonmedical personnel after they have received a modest amount of training.

An important point to consider when contemplating the collection of such clinical data is the sensitive nature of medical tests, particularly HIV tests. Clinical data should not be collected until clearance has been obtained from government agencies, particularly any agencies that deal with research on human subjects. Some organizations require that people who test positive for particular diseases be treated. Testing for HIV is a particularly sensitive issue; some organizations mandate that counseling be given to all people who receive HIV test results, even individuals who test negative for HIV.

**COGNITIVE TESTS.** Simple tests can be administered that measure the cognitive functioning of adults and the cognitive development of children. (See Chapter 7 on education for a brief discussion of cognitive tests.) Cognitive functioning measures for older adults can also be measured using tests of memory or tests of the ability to perform simple calculations or other mental tasks. These tests can be used to capture cognitive problems associated with aging.

**PROXY RESPONSE.** In some past LSMS surveys (for example, South Africa and Bulgaria), the health modules asked one individual in each household to answer questions about the health of all household members. Unfortunately, such data are likely to be very inaccurate. In particular, proxy responses can generate measurement error in the data, so it is extremely important that each adult in the household respond to questions about his or her own health status, utilization of health services, and health-related behavior. The only exception is that mothers should respond to questions about the health of any of their children who are too young to answer for themselves. The age at which older children can answer for themselves will vary from country to country, but a general range is somewhere between 10 and 15 years.

**CHANGES IN HEALTH STATUS.** It is sometimes desirable to measure changes in health status rather than the level of health status at only one point in time. A person's health status at a given point in time reflects past events that have occurred over an individual's entire lifetime, while a change in a person's health status between two points in time is primarily due to events that occurred between those two points in time. There are two possible ways to measure changes in health status. First,

respondents can be asked to give retrospective as well as current information about their health status. For example, as part of the self-reported activities of daily living, respondents could be asked to say how long they have been unable to perform the activity in question. However, retrospective data can be very inaccurate; a second, more accurate method is to collect data on a person's health status at two different points in time. This implies administering the same survey to the same households at two or more points in time—in other words, collecting panel data. See Chapter 23 for a general discussion of when to collect panel data as a part of an LSMS or similar multitopic survey.

**SUMMARY.** When household surveys are used to measure health status, survey designers should bear in mind the following points. First, health status is difficult to measure in household surveys, especially in multitopic surveys such as LSMS surveys, because it is usually not possible to use health professionals to give respondents thorough health examinations. Second, self-reported assessments of general health status can provide useful information, but they are also subject to bias. Third, self-reported assessments of current morbidity are not reliable, with the exception of reports of diarrhea. Fourth, self-reported activities of daily living are potentially very useful indicators of adult health status. Fifth, anthropometric measurements are reliable indicators of the health status of both children and adults, and have been used successfully in many developing countries. Sixth, if adequate resources are available it may be possible to perform a few simple medical tests; however, this is somewhat experimental and raises ethical issues concerning the use of human beings for research purposes. Finally, survey designers should ensure that each individual provides the interviewer with information about his or her own health; if proxy respondents are used, the results are likely to be unreliable.

*Using the Household Questionnaire to Measure Health-Related Behavior, the Use of Facilities, Health Expenditures, and Insurance Status*

As explained in the first section, collecting information on health status is only the starting point for analyzing health issues in developing countries. Data are also needed on health-related behavior, the use of health care services, health care expenditures, and insurance coverage. This subsection discusses how to obtain these kinds of information.

**HEALTH-RELATED BEHAVIOR.** A person's health status is closely related to his or her health behavior. The types of health behavior of most concern to policymakers, the health behaviors of the population, and the best ways to measure health behaviors all vary from country to country. Yet in most countries, two important kinds of health-related behavior are smoking and the consumption of alcoholic beverages. A relatively simple assessment of each can be done by asking respondents whether they have ever smoked or consumed alcoholic beverages, how old they were when they began, whether they are still doing so today, and how much they typically smoke or drink. Such data provide not only current information but also retrospective information on changes in these kinds of behavior over time, and on differences in behavior across population cohorts for a given age range.

Another type of health-related behavior that concerns policymakers is sexual practices. Because of the sensitive nature of this topic, collecting data on it is difficult. In particular, it is likely that respondents will underreport any risky behavior they engage in, such as failure to practice safe sex. The best way to collect data on sexual practices will vary across countries and across cultures. In many countries it may be impossible to collect such data as part of an LSMS-type household survey. Therefore, the draft health module does not include questions on sexual practices. For an example of surveys that do collect such data see the standard questionnaire used by the Demographic and Health Surveys (Macro International 1995a), which collects a small amount of data on sexual practices. Some Demographic and Health Surveys, including a 1994 survey in Tanzania (Macro International 1995b), have collected more detailed data on sexual practices. Demographic and Health Surveys collecting more detailed data on sexual practices have also been done in Brazil, Burkina Faso, the Central African Republic, Côte d'Ivoire, Haiti, Kenya, Uganda, and Zimbabwe.

Household-level behavior that affects the epidemiological environment of the household includes cooking practices (the type of fuel used, the extent of ventilation for the smoke, and whether there is a separate room for cooking), waste disposal, sanitation practices, and the sources of water used by the household. LSMS surveys typically ask for this kind of information in the housing module. For a detailed description of how to collect such data, see Chapter

12 on housing and Chapter 14 on environmental issues.

Other types of health-related behavior include the use of preventive medical care, diet (including infant feeding practices), exercise, and the use of seatbelts. Information on preventive care should be collected together with data on utilization of health care facilities. Data on infant feeding should be collected in the fertility module (discussed in Chapter 15). Diet information for other household members is very difficult to measure in household surveys and thus probably should not be collected in most LSMS surveys. For further discussion of this see Chapter 5, which discusses the collection of food consumption data. Finally, it is relatively simple to ask questions on the use of seatbelts (and the use of helmets among motorcyclists), on work-related physical activity, and on personal exercise. An example of this is the LSMS survey done in Brazil in 1996, which had seven questions on personal exercise.

**UTILIZATION AND EXPENDITURES.** Most previous LSMS surveys have collected incomplete data on utilization rates and expenditures. Individuals have typically been asked whether, in the month preceding the survey, they have seen a health care provider to be treated for an illness and, if so, how much they spent on that visit. However, individuals, particularly individuals with serious illnesses, may make several visits to receive health care over one month and may obtain treatment from more than one provider. Since most previous LSMS surveys have not been designed to capture this comprehensive information, the data from such surveys probably underestimate both utilization rates and expenditures. This is why most studies of the demand for health care that use LSMS data are studies of provider choice rather than studies of expenditures or general studies of the demand for health care services.

Data on all health care consultations and expenditures are required for many kinds of policy analysis. Analysts need data on all visits to medical facilities to see which socioeconomic groups use which facilities, and they need data on all expenditures to obtain an accurate measurement of the costs to households of obtaining health care. Ideally, data should be collected by the level of care (primary, secondary, or tertiary), by the type of provider (public, private, or traditional), by the purpose of the visit (preventative, curative, or prenatal care), and by the kind of services received. To

avoid problems of recall bias, questions on any outpatient care received should be limited to the previous 30 days. Since inpatient care is less common and easier to remember, the recall period for questions about inpatient care can be the previous 12 months. Expenditures should include not only fees but also any other expenses incurred by the respondents, such as purchases of medicine and transportation costs. It is also useful to ask about the amount of time spent in obtaining the care (including travel time and any time spent waiting to see the provider), since this is another cost of using these services. Finally, for households or individuals that have health insurance, it is important to distinguish between charges paid for or reimbursed by the insurance and charges paid for by the respondent.

As mentioned in the first section of this chapter, one of the most important uses of data on the utilization of health care services is to study the incidence of government health subsidies. This is done in two steps. The first step is to measure the unit subsidy provided by the government for each type of health care service (such as inpatient care, lab tests, primary care, prenatal care, and immunizations). The unit subsidy is the incremental cost of providing the service minus the fee charged. Estimates of the incremental cost can be obtained from cost function estimates based on facility data, either obtained from the ministry of health or collected in a survey of health care providers.<sup>3</sup> A recent example of such estimates, done in the Philippines, is given by Alba (1998).

The second step in this benefit incidence analysis is to calculate the amount of the overall subsidy received by different population groups. This calculation can be done by multiplying the unit subsidy by the utilization rate of that service by each group in the population. Thus benefit incidence analysis enables policymakers to know how public subsidies for specific types of health care services are distributed among the general population by geographical location, socioeconomic status, education, age, and sex.

**INSURANCE AND EMPLOYEE HEALTH BENEFITS.** Analysts can use data on insurance and employer-provided health benefits to explore how these benefits affect health status, health behavior, and the utilization of health care services. The structure of health insurance and of employer-provided health benefits differs from country to country. At minimum, respondents must be

asked whether they are covered by some sort of insurance or health benefits. However, in many countries that have health insurance and employer health benefits, the nature of both varies widely according to the type of insurance policy or the type of employer. Thus analysts also need data on the benefit structure to be able to measure how health insurance or employer health benefits affect an individual's health status, health behavior, and utilization of health care services. For example, the fact that an individual is covered by health insurance is unlikely to cause him or her to make greater use of outpatient services if the insurance policy covers only inpatient care. Thus, for health insurance, data should be collected on the services covered by the insurance, including whether the policy covers private as well as publicly provided services. People may also base their health care decisions on their policy's structure of deductibles, co-insurance (that is, required copayments), and benefit ceilings. Information is also needed on which household members are covered by an individual's health insurance; some policies cover only the individual while others cover all (or some) of the members of the individual's family. Two studies that examine the impact of insurance are Brook and others (1983) and Gertler and Sturm (1997). Similarly, for employer-provided health benefits, data are needed on the number of sick days allowed, the wages received for those sick days, and analogous information for maternity benefits.<sup>4</sup>

There is another aspect of insurance and employer health benefits that policymakers should consider. Doing a full analysis of benefits requires complete information on the financial risk associated with bouts of illness. This in turn requires data on all of the expenditures made by an individual for treatment of a specific episode of illness—in other words, the expenses associated with all consultations with health care providers from the onset of the illness until the illness is cured. This is a complex task because expenditures on health care can vary greatly during the course of treatment. Also, in some cases the utilization of health care for a specific illness may have started before, or may continue after, the 30-day recall period recommended for LSMS-type surveys, especially in the case of individuals who have severe illnesses. Unfortunately, it is practically impossible to collect this type of information by asking these respondents to remember details of their treatment over a long period of time, so it is probably impossible to make a full calculation of

the benefits of insurance using LSMS-type surveys.<sup>5</sup> For an idea of the difficulty involved in doing such a study, see Gilleskie (1998).

**ACCESS.** To understand the determinants of the use of health care services, analysts need data on the extent to which the population has access to health care facilities. This can be done by asking households about nearby facilities and perhaps about more distant facilities that they may use regularly, and then gathering data on those facilities using a facility questionnaire that is part of the overall survey or using an existing source of data unrelated to the survey (such as data collected by the ministry of health). The main problem with this method is that it omits any nearby providers whose existence is unknown to the sampled households. This problem leads to the more general issue of sampling health care providers (discussed further below in a separate subsection).

**DATA FROM OTHER MODULES.** It is important to stress that doing a causal analysis of how health status affects other socioeconomic outcomes (as well as how non-health factors affect health status) requires not only the health status variables discussed above but also data on other household characteristics and behavior.

Consider first the impact of health status on other socioeconomic outcomes. Common outcomes of interest to policymakers include income (especially labor income), education, labor force participation, migration, and fertility. Data on these outcomes are collected in the questionnaire modules discussed in the analogous chapters of this book. For example, data on wages, which indicate worker productivity, are found in all versions of the employment module introduced in Chapter 9 (and presented in Volume 3).

Turn now to the effect of other characteristics on health outcomes. Possible causal factors are housing characteristics (such as source of water, type of toilet, and method of garbage disposal), the education levels of adults, labor force outcomes, characteristics of the local environment, household income, fertility history, and perhaps access to credit. Which income measure is most useful will depend on the particular issue being analyzed. In some cases the best measure will be some type of permanent income that is less subject to year-to-year fluctuations than annual income. In other cases it may be more desirable to measure only nonlabor income. If several years of data are available, permanent

income can be approximated by averaging income over several years, and in some cases permanent income can be proxied by per capita consumption. See Chapter 17 for a discussion of the issues involved in measuring income and Chapter 5 for a discussion of household consumption data, which can be used as an indicator of permanent income. Chapter 7 discusses the collection of data on education outcomes. Finally, data on prices are needed to draw causal inferences about how all of these characteristics affect health outcomes, as explained in Appendix 8.1. Price data are collected in the price questionnaire, which is discussed in Chapter 13.

Finally, a word of caution. Using household survey data to estimate causal relationships must be done with great care; many pitfalls and complications are involved. In general, simple techniques such as ordinary least squares regressions are likely to produce biased results. This is discussed further below.

#### *Prices, Quality, and Public Health Programs: The Facility, Price, and Community Questionnaires*

Policy issues related to the price of health services, the quality of health services, and public health programs can be analyzed by combining data on health status, health behavior, utilization of health services, and expenditures on health care (all of which are collected in the household questionnaire) with data on local health facilities and programs. The need for data on health facilities and programs moves the discussion from collection of household data to collection of data that characterize the community in which the households live. Although this type of information could be collected at the household level, doing so would be inefficient (because much of the information does not vary across households living in the same community) and probably inaccurate (because many households may not be familiar with some of the information being sought).<sup>6</sup> A better approach is to collect this type of information at the community level. Data on the quality and prices of health services can be obtained from a sample of local health facilities, including both private and government-operated facilities. Information on public health programs can be collected in the community questionnaire. Finally, the prices of any medicines that can be purchased from pharmacies or other vendors can be collected using the price questionnaire. This subsection discusses each of these types of information in turn.

**PRICES.** In order for policymakers to understand the financial pressures of health care on households, they must have information on the full range of costs that households face when using health care services. The most obvious costs are payments to health care providers, which are determined by the prices charged by health care facilities. This information can be obtained from health facility questionnaires, yet several issues arise when calculating the price of health care services. First, visits to health care facilities vary widely in terms of what takes place during those visits. For example, a single visit may involve both diagnostic and treatment services.

One common approach to measuring the prices of health care services is to calculate the mean or median expenditures per visit, but this approach does not account for variation in the types of visits. A better approach is to collect prices from health care facilities for specific types of visits. One use of such data is to construct a price index of health care costs that summarizes the average “price” charged by each facility. Two types of data are required for these price indices: data on the prices of the services provided and data on average household expenditures on each service. The data on health care expenditures can be used to calculate weights for the health care price index, just as household expenditure data on all goods and services can be used to calculate weights for general price indices.

A second issue regarding the prices charged by health care providers is that prices may be higher for health facilities that provide higher quality services. To understand the pure effect of prices on household choices and health outcomes (in other words, the effect of a change in price for a given level of quality), it is necessary to remove any variation in prices that simply reflects variation in quality. If this is not done, estimates of the effects of prices on health outcomes and on demand for health care are likely to be biased. However, estimating such quality-adjusted prices is not an easy task. It can be done only at the data analysis stage, not at the data collection stage.

There are several ways to adjust prices for quality differences at the data analysis stage, all of which use facility data on service quality. (Such data are discussed further below.) One way is to specify quality as a provider fixed effect, which makes it possible to purge the impact of quality from the prices of specific services. (See Deaton 1988 for a specific example; see Chapter 23 for a general discussion.)

An alternative approach, of which Goldman and Grossman (1978) offer an early example, is to use hedonic pricing models to control for differences in quality from provider to provider that are reflected in the prices that the providers charge. This approach involves estimating how a variety of dimensions of quality affect the prices charged by a sample of providers for a range of different services. The difference between the actual price and the predicted price from this regression is the quality-adjusted price. If one assumes that the predicted price from the regression reflects all differences in quality, then the difference between that price and the actual price reflects differences in price that are not due to quality differences.

In general it is possible to estimate the impact of prices on health outcomes only if data are collected from individual health care providers using a facility questionnaire, although there are some exceptions (for example, if all health facilities are operated by the government and there is no variation in price among them). The same is also true for analyzing how the quality of services affects health outcomes; such analysis cannot be done unless data are collected from health care facilities.

A third issue is that fees paid to health care providers are not the only costs of obtaining health care, even after including unofficial charges such as tips and bribes. There are also the financial and time (opportunity) costs of traveling to and from a facility and the cost of the time spent obtaining treatment. The financial costs are simply a person’s expenditures on transportation. The time cost includes the travel time to a facility plus the time spent at the facility, multiplied by the value of the time of the person treated (and of anyone accompanying him or her). To compute these time costs, analysts need data on respondents’ travel times and waiting times, as well as their wage rates.

Travel times are best collected from each household, although one could collect this information at the community level if all households in the sampling unit live very close to each other.<sup>7</sup> Wage rates can be obtained from the employment module of the household questionnaire (see Chapter 9). It is useful to collect data on waiting time in a facility questionnaire as well as in the household questionnaire. Collecting data on waiting time in the facility questionnaire is useful for calculating quality-adjusted prices and also avoids inaccurate average waiting times based on only one or two

households who actually used the facility. Yet asking about waiting time in the household questionnaire can be useful for estimating simple averages across different types of health facilities. The effects of any insurance on prices paid can be captured by including information on deductibles and copayments when calculating prices. This insurance information must be collected in the health module of the household questionnaire.

A final cost of health care is the cost of purchasing medicines from outlets that are not health care providers, such as pharmacies and small vendors. The best place to collect this information is in the price questionnaire (see Chapter 13 for a full description), which can be used to collect price data for many kinds of commonly purchased medicines.

**QUALITY OF CARE.** The data necessary to evaluate how the quality of care affects household behavior and health outcomes should be collected in the facility questionnaire. The best approach is to ask providers what actions they take under various circumstances and what resources they use in providing health care services. Some households may be able to provide this information, but others may not or may provide inaccurate information.

The quality of the care provided by each facility or provider depends on both the structure of care and the process of care. The structure of care is the quantities and types of inputs (such as equipment, personnel, and medicine) used by the provider in providing its health care services. The process of care is the way in which services are provided during a patient's visit, including the way in which professionals diagnose and treat patients with specific health needs. For example, in the Jamaican LSMS survey conducted in November 1989 process measures were collected from both facility and household questionnaires to assess the quality of prenatal care.

There is evidence that the process of care varies substantially among developing countries. Peabody (1996) noted that inaccurate diagnoses and inappropriate treatment are common in rural Vietnam; diarrheal disease is often inappropriately treated with antibiotics rather than with oral rehydration therapy. Using data from Jamaica, Peabody, Gertler, and Leibowitz (1998) found that a better process of care (measured by how actual care compared to optimal diagnosis, treatment, and advice protocols) was associated with a 500-gram increase in the birthweights of

children whose mothers received that care—after controlling for variation in households' socioeconomic characteristics and in risk factors.

Information on the structure of care can be obtained by asking about equipment, personnel, and the availability of different kinds of medicines. Information on the process of care can be obtained by asking about the protocols followed for different types of commonly provided health services. Several examples are in the draft facility questionnaire introduced in the third section of this chapter (and presented in Volume 3 of this book). For a more detailed discussion see Peabody, Gertler, and Leibowitz (1998) and Peabody and others (forthcoming).

**PROVIDING PUBLIC HEALTH SERVICES TO THE GENERAL POPULATION.** Public health services and programs are generally outreach programs and thus are not necessarily associated with particular health care facilities. This implies that data on all public health services and programs available in a given community, such as immunization programs and information campaigns, should be collected in the community questionnaire. Data should also be collected in the community questionnaire on the existence (and quality) of public water supply and garbage collection services. In addition, it is useful to collect information on air and water pollution, although it is difficult to collect this accurately. The community questionnaire introduced in Chapter 13 collects all of this information in detail, except for data on air and water pollution, which can be gathered in the relevant submodules presented in Chapter 14. Finally, the household questionnaire should include explicit questions on who in the household participated in these public health programs—for example, who was immunized, received a pamphlet, or was otherwise affected by an information campaign.

**REGULATION OF PRIVATELY PROVIDED HEALTH CARE SERVICES.** As explained in the first section of this chapter, the main role of the government regarding privately provided health care is the regulation of private health care facilities. Data from LSMS surveys can be used to measure the extent to which these facilities comply with the relevant regulations, especially if a health care facility questionnaire is included as part of such a survey. Specifically, data from a facility questionnaire can show whether the staff at the facility have the required train-

ing and licenses, whether the facility's equipment meets certain standards (for example, whether each piece is clean and in working order), whether proper procedures are being followed for specific types of examinations and treatment, and whether regulations regarding drugs and prices are being followed.

Additional information on compliance can be gathered from the household questionnaire. Such data can be used as an independent check of drug and price regulations and of whether certain procedures are being followed. (This is particularly useful if health facilities provide misleading information about their compliance with government regulations.) It may also be possible to check some regulations about insurance using data from the household questionnaire.

A second general use of data from the health facility questionnaire—and, to a lesser extent, from the household questionnaire—is to provide a picture of the private health care system. For example, these data may reveal deficiencies in the provision of private health care, and they may also suggest to policymakers specific regulatory actions to resolve these problems. Such information can be useful to policymakers when they are considering new regulations on private facilities.

A third use of data from private health care providers is to examine how regulations affect these providers' services and even their patients' health outcomes. At minimum, the extent to which private facilities comply can be assessed, as discussed above. In addition, if data indicate that some providers comply while others do not, it may be possible to use the variation in compliance to estimate the impact of compliance on health outcomes. However, such estimation is not straightforward and could lead to biased results. Finally, if panel data are collected and new regulations are put in place during the time spanned by the surveys, it may be possible to estimate the impact of regulations on the phenomena that are being regulated and perhaps even on health care outcomes. Yet here too the estimation problems are considerable, and analysts should be cautious about drawing causal inferences. For further discussion of estimation problems see Appendix 8.1 and Chapter 23 on panel data.

Investigating the behavior of households and of private providers is a fourth way to use data on private health care facilities. First, policymakers need to know how changes in the public provision of health care services affect the use of private health care providers. For example, if increasing prices at public facilities

reduces the extent to which these facilities are used, one would like to know how much of this is due to people switching to private providers and how much is due to people receiving no professional health care at all. Second, private providers may alter their behavior in response to public health care policies. In theory, this could be analyzed using data on private providers—ideally using estimates of the supply curve of different kinds of privately provided health care. Yet such analysis would be fairly experimental. It is difficult to say whether such analysis is feasible using data from an LSMS-type household survey.

Fifth and finally, if detailed cost data are collected from private health care facilities, these data can be used to estimate the cost of providing different services. Such information might be used to assess the efficiency of the public sector, as well as to undertake cost-benefit analysis. However, this is also a new area of research, one for which there may be many unforeseen analytical problems.

### *Using Data for Policy Analysis*

The discussion up to this point has said little about how to use these data to analyze health policies and programs. A detailed discussion would require a separate chapter, if not an entire book. Yet a brief discussion can highlight some fundamental points that survey designers should bear in mind when designing the health module for an LSMS survey. This subsection briefly reviews some methodological issues that arise in policy analysis in the health sector.

In general, policy questions can be divided into two categories:

- Questions about overall levels of outcomes and the distribution of those outcomes across subgroups of the population.
  - Questions about the effects of policies on outcomes, both in the aggregate and within subgroups.
- Table 8.1 lists 37 policy questions. Of these, questions 1–14, 18, 25, and 28–33 concern the overall level and distributions of outcomes, while questions 15–17, 19–24, 26, 27, and 34–37 pertain to the effects of policies on outcomes. The research methods needed to answer the first set of questions are straightforward. The questions can be addressed by describing the current situation. All that is needed is descriptive information, such as the means of the outcome variables both in the aggregate and by population subgroup. Analyzing the second set of policy questions raises dif-

**Table 8.1 Policy Questions and Data Sources**

Policy question	Minimal data needs from a household survey	Expanded data needs from a household survey	Sources other than a household survey
1. Who is getting the most serious diseases?	Diarrhea only: short 4–8, standard A6–A10	Vital registration data	
2. Who is unable to perform their usual activities because of poor health?	Short 1, standard A3		
3. What is the level and distribution of health in terms of physical functioning within the population?	Standard A11–A29	Expanded G1–G11	
4. What is the level and distribution of health in terms of cognitive functioning within the population?	Adult: expanded H1–H32; child: Education module		
5. What is the level and distribution of health in terms of mental health within the population?	Expanded A30–A47		
6. What is the level, cause, and distribution of adult mortality in the population?			Vital registration data
7. What is the level, cause and distribution of child mortality in the population?	Fertility module		Vital registration data
8. Who is engaging in individual health-related behavior of particular concern? What is the degree of participation in such behavior?	Standard B2–B23	Fertility module (breastfeeding and weaning)	
9. Which households are engaging in sanitation, waste disposal, and water supply practices that adversely affect the health of their members?	Housing module		
10. Who is obtaining health care by purpose, provider type, and level of care?	Short 9–35 (only provider type), standard E1–E54 (more detail, but no data on purpose)	Expanded E1–E90 (very detailed, including by purpose)	
11. What providers and services do households have access to?	Community questionnaire	Standard F1–F8	
12. What are private expenditures on care by purpose, provider type, and level of care?	Short 11–37 (brief), standard E3–E58 (more detailed)	Expanded E8–E94 (very detailed)	Special interviews or calendars for severe illness
13. What is the utilization of each subsidized service, in the aggregate and by subgroup?	Short 9–35 (brief), standard E1–E54 (more detailed)	Expanded E1–E90 (very detailed)	
14. What is the unit subsidy for providing each service?	Expanded E8–E90, facility questionnaire B4, B9, B11		Estimates of incremental cost from cost function
15.* How does a change in public education, health promotion activities, and other public health activities affect health?	Measures of health: short 1, 4–7, standard A3, A6–A9, A11–A29; anthropometric data: Anthropometry module; education and activities: Community module	Additional health measures: expanded A30–A47, G1–G11, H1–H32	
16.* What is the effect of the quality of the air or water supply on health?	Measures of health: short 1, 4–7, standard A3, A6–A9, A11–A29; anthropometric data: Anthropometry module; environment data: Environment modules	Additional health measures: expanded A30–A47, G1–G11, H1–H32.	
17.* What are the effects of government health promotion activities on individual health-related behavior and on knowledge of the health risks associated with certain behavior?	Health behavior: standard B2–B23; government activities: Community module		Other information on health promotion activities
18. What is the price that households pay for care by type of service, level of care, and provider type?	Standard E3–E58 (only by provider type), facility questionnaire B4, B9, B11	Expanded E8–E94 (full detail)	

**Table 8.1 Policy Questions and Data Sources (continued)**

Policy question	Minimal data needs from a household survey	Expanded data needs from a household survey	Sources other than a household survey
19.* What is the effect of a change in user fees for a service on utilization of that service, utilization of other services, provider choice, household expenditures on health care, and government revenues from user fees by service?	Standard E1–E54, facility questionnaire B4, B9, B11	Expanded E1–E90 (more detail on utilization and expenditures, can also use to impute prices)	
20.* How does a change in user fees affect health?	Measures of health: short 1, 4–7, standard A3, A6–A9, A11–A29; anthropometric data: Anthropometry module; fees: facility questionnaire B4, B9, B11	Additional health measures: expanded A30–A47, G1–G11, H1–H32; can also use E1–E91 to impute prices	User fees
21.* What is the effect of program location on utilization of services, provider choice, household expenditures on health care, and government revenues by service?	Standard E1–E54; program location: Community module; standard F1–F8	Expanded E1–E90 (more detail on utilization and expenditures)	
22.* How does a change in program location affect health?	Measures of health: short 1, 4–7, standard A3, A6–A9, A11–A29; anthropometric data: Anthropometry module; program location: Community module; standard F1–F8	Additional health measures: expanded A30–A47, G1–G11, H1–H32	
23.* How do changes in the availability of government services affect household health-related behavior?	Health behavior: standard B2–B23; government services: Community module; standard F1–F8		
24.* What is the effect of a change in alcohol or tobacco taxes on consumption of those goods, on household expenditures on these goods, and on government revenues from the taxes?	Health behavior: standard B2–B16; prices: Community module		
25. What is the quality of care available to households by type of service, level of care and provider type?	Standard F1–F8, facility questionnaire		
26.* What is the effect of a change in the quality of a service on utilization of that service, provider choice, household expenditures on health care, and government revenues by service?	Standard E1–E54; service quality: facility questionnaire (Parts A–F)	Expanded E1–E90 (more detail on utilization and expenditures), facility questionnaire (Parts G–K)	
27.* How does a change in the quality of health care affect health?	Measures of health: short 1, 4–7, standard A3, A6–A9, A11–A29; anthropometric data: Anthropometry module; service quality: facility questionnaire (Parts A–F)	Additional health measures: expanded A30–A47, G1–G11, H1–H32; facility questionnaire (Parts G–K)	
28. What is the quality of health care available from private sector providers?	Facility questionnaire (Parts A–F)	Facility questionnaire (Parts G–K)	
29. Are private sector providers following government regulations regarding service prices, the quality of services, and the use of pharmaceuticals?	Facility questionnaire (Parts A–F); some price data also from standard E3–E58	Facility questionnaire (Parts G–K); price data also from Expanded E8–E94	
30. Are private providers of insurance following government regulations?	Standard D1–D17		Special study on insurance providers
31. What is the impact of government regulations on the quality of care in the private sector?	Facility questionnaire (Parts A–F; panel data)	Facility questionnaire (Parts G–K; panel data)	Information on government regulations
32. Who is covered by health insurance? What are the benefits, services covered, copayments, deductibles, and benefit caps?	Short 38, 39 (coverage only)	Standard D1–D17 (detail on costs, benefits, and so on)	

(Table continues on next page.)

**Table 8.1 Policy Questions and Data Sources (continued)**

Policy question	Minimal data needs from a household survey	Expanded data needs from a household survey	Sources other than a household survey
33. Who has health benefits at the place of employment? What are these health benefits?	Employment module		
34.* How does the presence of insurance coverage affect health care utilization patterns, expenditures on health care, and government revenues from user fees?	Standard E1–E58; for insurance: short 38, 39	Expanded E1–E94 (more detail on utilization and expenditures); for insurance: standard D1–D17 (detail on costs, benefits)	
35.* What is the effect of health on worker and farm productivity?	Measures of health: short 1, 4–7, standard A3, A6–A9, A11–A29; anthropometric data: Anthropometry module; for worker and farm productivity: Employment module, Household enterprise module, Agriculture module	Additional health measures: expanded A30–A47, G1–G11, H1–H32	
36.* What is the effect of health on education and cognitive outcomes?	Measures of health: short 1, 4–7, standard A3, A6–A9, A11–A29; anthropometric data: Anthropometry module; education outcomes: Education module	Additional health measures: Expanded A30–A47, G1–G11, H1–H32	
37.* What is the effect of health on consumption and savings?	Measures of health: short 1, 4–7, standard A3, A6–A9, A11–A29; anthropometric data: Anthropometry module; savings and consumption: Savings module, Consumption module	Additional health measures: expanded A30–A47, G1–G11, H1–H32	

Note: Questions marked with a (\*) require causal analysis. The numbers in the second and third columns indicate question numbers from the different versions of the health module in the household questionnaire. For the standard and expanded modules the letters refer to the "Part" of the module. For example, "standard A6" refers to Question 6 of Part A of the standard health module.

Source: Authors' summary.

difficult estimation issues because it requires estimation of causal relationships—assessing how specific policies affect complicated behavioral choices.

This subsection briefly discusses some methodological issues that arise in answering the second set of questions. A detailed (and more technical) discussion of these methodological issues can be found in Appendix 8.1. Analysts may want to know how a specific health program affects a specific health outcome, such as the incidence of diarrhea. One simple approach is to compare the mean incidence of diarrhea among communities that benefit from the program with the mean incidence in communities that do not benefit from the program. Unfortunately, this approach cannot be used to determine the causal impact of the program. Suppose a simple analysis of means indicates that communities benefiting from the health program have a lower incidence of diarrhea. This may not be due to the program, because communities that benefit from the program may also have higher income levels, and higher income levels may be responsible for all or part of the lower incidence of diarrhea. Thus, while a causal relationship may exist, it can easily be obscured by other factors that have not

been controlled for. Another example is user fees. Suppose that user fees are higher in areas where households' incomes are relatively high. In this case, an analysis of the relationship between user fees and the utilization of health care services that does not control for households' income levels may show that higher user fees are associated with higher levels of health care utilization, because higher income leads to greater utilization. The basic statistical tool for estimating causal relationships is multiple regression analysis, which controls for the effects of all variables that can potentially affect the health outcome being studied.

For the analysis of health issues, the two most important relationships to estimate using multiple regression analysis are health input demand equations and health demand equations.<sup>8</sup> Estimates of health input demand equations measure the determinants of health care utilization, expenditures on health care, and health-related behavior. Each of these health-related outcomes is determined by a set of causal variables, namely prices (both for health services and for goods unrelated to health), prevailing wage rates, household income, other household characteristics (such as the education levels of household members),

and a variety of community characteristics, including the characteristics of locally available health services. In addition, households vary in their “tastes” for health and in the “innate healthiness” of individual household members.

The data required to use multiple regression analysis to estimate a health input demand equation are the health input of interest (such as the use of health care services, expenditures on health care, or a specific health-related behavior) and the entire set of variables that determine the demand for that input. While it is possible (though not necessarily easy) to collect data on prices, wages, income, and many other household and community characteristics, it is much harder to measure tastes for health and the innate healthiness of individuals. This inability to observe some of the variables that determine households’ demand for these health inputs leads to a serious estimation problem: omitted variable bias (discussed further below). In fact, the two examples above (diarrhea, user fees) on the problem of using simple comparisons of means can be thought of as cases of omitted variable bias, in which the omitted variable is household income.

Health input demand equations can be used to estimate how specific policies or programs, such as public health programs and price changes, affect health care utilization, provider choice, and health-related behavior. Of course, the price that really matters is the effective price paid by the consumer. This means that in addition to the data on the price charged by the health care facility, data are needed on taxes, travel expenses, waiting times, wage rates, and insurance copayments. The effects of travel and waiting times can be accounted for either by including time costs in the calculation of the price or by including distance in the equation as an additional determining variable. Estimates of the effects of specific policy changes on government revenues from taxes and user fees can also be calculated from estimates of how prices affect utilization rates.

What about the impact of government policies on health status? Clearly, health status is strongly affected by health care utilization, expenditures on health care, and health-related behavior. Thus the same variables that determine these demands for health inputs also determine health. This relation is a health demand equation (also known as a health demand function). Since current health can depend on past health, it is

often more useful to estimate changes in health status as determined by prices, wages, household income, and other household and community characteristics. For example, a child’s nutritional status at 12 months of age can be thought of as being determined by his or her nutritional status at 6 months of age and the health inputs that he or she has received during the intervening 6 months.

Perhaps the most important lesson for survey designers from this discussion of methodological issues is that if they want to analyze the causes of health status and related activities (health care utilization, health care expenditures, and health-related behavior), it is necessary to collect data on all of the characteristics of households and communities that determine these health outcomes. Not having all of these data can lead to omitted variable bias. The magnitude of the bias will depend on the correlation of the omitted variable and the included variables and on the magnitude of the omitted variable’s true effect on the outcome.

Omitted variable bias can arise for several reasons. First, it can arise if one of the observable determinants of the health outcome is omitted, as in the above examples in which household income was omitted from studies of diarrhea and user fees. Second, omitted variable bias can also arise because some of the determining variables, such as tastes and innate healthiness, are almost impossible to measure. A third source of omitted variable bias arises in the case of a dynamic model, as explained in Appendix 8.1. For further discussion of omitted variable bias and how to avoid or at least minimize it, see Appendix 8.1 and Chapter 26.

### *Sampling Issues*

Two issues need to be considered with respect to sampling. First, LSMS surveys should continue their practice of not subsampling individuals within a household. This would yield samples that would be too small to draw reliable inferences about the utilization of health care. For example, in the 1987–88 LSMS survey in Ghana, 5,746 individuals (39 percent of the sample) reported being ill or injured during the previous four weeks, and 2,398 (17 percent of the sample) had received medical attention. Disaggregating these visits by type of provider shows that 1,106 were treated by doctors, 292 by nurses, 654 by medical assistants, 84 by midwives, 70 by pharmacists, 109 by healers, and fewer than 100 by other providers. These Ghanaian households had an average of 4.68 members, so randomly

sampling two individuals per household would have yielded numbers less than half of those reported above. Such samples are too small to generate reliable estimates of expenditures and utilization rates, particularly by provider. The sample sizes for rare illnesses and their associated treatments would be even more miniscule.

A second issue regarding sampling is the selection of the health care facilities to which to administer the facility questionnaire. In rural areas there may be few facilities to visit, so all nearby facilities (and possibly even some that are far away) should be visited. In contrast, in large urban areas the number of facilities that households could potentially visit is so large that it is not feasible to visit them all or even most of them. In such cases survey designers must choose a sample of facilities. In an LSMS-type survey the facilities of interest are those that are available to the households in the sample. Thus the general principle is to ensure a good deal of overlap between the facilities interviewed and the facilities that respondents know about and use. The best method for choosing such facilities, discussed in more detail in Chapter 13, is to compile a list of facilities based on those mentioned most often in responses given in the household questionnaire. If the number of facilities on the cumulative list is small, or if the survey budget is large, all facilities on the list can be included in the sample. Otherwise, a sample of facilities should be drawn from the list. The sample of facilities can be selected either randomly or according to some other criterion, such as using a probability proportional to the number of times they are mentioned by household respondents. Once selected, the list of facilities and their associated code numbers is recorded in the community questionnaire.

#### *Links Between Policy Issues and Data Needs*

This subsection links the policy issues discussed in the first section of the chapter to their specific data requirements, with reference to the draft modules introduced in the third section of the chapter. For convenience, this information is summarized in Table 8.1, which identifies the minimal data needs, expanded data needs, and data sources other than household surveys required to answer the policy questions.

There are two differences between the minimal and expanded data needs. The first is precision. Minimal data needs are based on relatively few questions and thus do not measure the particular phenomena of interest as accurately as do the expanded data needs, which are

based on more, and more detailed, questions. The second difference is that expanded data needs measure more dimensions of the phenomena of interest. For example, expanded data needs measure health outcomes that are not measured by the minimal data needs, such as mental health and cognitive functioning.

Another feature of Table 8.1 is that it distinguishes between policy questions that require causal analysis and policy questions that do not. Questions that require causal analysis are marked with an asterisk next to the number pertaining to the question. As pointed out above, the data requirements for causal analysis are quite large. In particular, all the determining variables that belong in the demand equations for health or health inputs, such as prices, wages, household income, and other household and community characteristics, are needed. There are some circumstances in which reliable estimates can be obtained without some of these variables. The issues involved in deciding which variables must be included are complex and beyond the scope of this chapter. However, a brief discussion and some useful references are provided in Appendix 8.1.

#### **ASSESSING HEALTH STATUS AND MEASURING ASSOCIATED BEHAVIOR.**

Questions 1–9 in Table 8.1 refer to policy issues associated with assessing health status and measuring associated behavior. Because of measurement problems associated with self-reported morbidity data, LSMS surveys are, in general, not very useful for gathering information on the incidence of specific diseases. Yet there are cases in which self-reported symptom data may be helpful, and carefully worded questions can provide useful measures for policy analysis. An important example is diarrhea; several questions about the incidence of diarrhea are contained in all three versions of the draft health module (for the standard and expanded module these questions are in Part A: Self-Reported Health Status). In some countries additional data on symptoms may be useful in assessing morbidity with self-reported symptoms, yet better data on the incidence of most diseases can be obtained from facility data and from official government statistics on deaths (which often include the cause of death).

In addition to data on morbidity, policymakers often want data on both adult and child mortality. Because adult mortality is a relatively infrequent event, household surveys with samples of 2,000–5,000

households—the recommended size for an LSMS-type survey—will not yield sufficiently precise measures of adult mortality for use in policy analysis. Official government statistics are a better source for this information. On the other hand, because child mortality is more frequent, some useful estimates can be obtained from data collected in the fertility module (see Chapter 15).

The only information on health outcomes in the short version of the draft health module is for diarrhea among young children. The standard and expanded versions collect both diarrhea information and self-reported activities of daily living in Part A (Self-Reported Health Status). The expanded version also collects mental health information in Part A. In addition, the expanded version collects data on observed activities of daily living in Part G and on cognitive functioning in Part H. Both the standard and the expanded versions collect information on health behavior—specifically, behavior related to the consumption of alcohol and tobacco—in Part B (Health-Related Behaviors). Because the behavior of concern will tend to vary by country and the wording of the questions will be culturally specific, Part B will probably need to be modified to reflect the health priorities in each country. Other kinds of behavior that might be addressed in Part B include exercise, the use of seat belts, and safe sex practices. Information on infant and child feeding practices can be collected in the fertility module (see Chapter 15). Information on sexual practices is much more difficult and thus is not attempted in this module. Finally, some information on knowledge of sexually transmitted diseases is collected in Part B of the standard and expanded modules. Similar information could be collected on knowledge of the risk of cancer from smoking or the risk of liver disease from consumption of alcohol.

**EQUITY ISSUES.** Questions 10–14 in Table 8.1 address equity issues. A key equity concern is the utilization of health services and associated expenditures. The short version of the health module collects basic information on utilization and expenditure in the previous four weeks but does not disaggregate expenditures by type (such as clinic fees, purchases of medicine, and transportation). The standard module collects significantly more information on utilization and expenditures in Part E (Health Care Utilization and Expenditures). Part E of the expanded version of the

health module collects very detailed information on utilization and expenditures in the previous four weeks.

Equity is also affected by differences in the availability of health care services and the extent to which people know that these services are available. The community questionnaire presented in this book asks a group of community representatives and leaders about what health care providers exist in their community. In addition, in Part F (Health Provider Knowledge) of the standard and expanded versions of the health module, each household is asked to name the closest health care providers of different types of which they are aware.

Finally, studying the distribution of government health care expenditures across different population groups requires not only data on the utilization of each service (which are gathered in Part E of the module) but also data on the unit cost of providing each service. Rough estimates of the unit costs for some services can be obtained from the facility questionnaire. More precise measures of unit costs require econometric estimates from a cost function for the provision of health care, which entails collecting a substantial amount of additional facility-level data.

**PUBLIC HEALTH PROGRAMS.** Questions 15–17 in Table 8.1 address three policy issues regarding public health programs. The first two questions concern how public health policies affect health outcomes. At a minimum, answering these questions requires some measure of health and a measure of the policy variable that varies over the sample. The data needed for assessing health status were discussed above. The policy variables are collected primarily in the community questionnaire, and perhaps also in the facility and price questionnaires.

Question 17 in Table 8.1 concerns health behavior and health knowledge. Data on specific types of behavior and knowledge of the health risks are found in Part B of both standard and expanded versions of the health module. Data on government health promotion activities should be collected in the community questionnaire.

**PRICING POLICIES FOR HEALTH CARE SERVICES.** Questions 18–24 in Table 8.1 address health care pricing policies. In general, the short version of the health module contains only a small amount of data on these issues, while both the standard and expanded modules

collect much more information. There are several options for measuring health care prices. Community-level prices for commonly available medicines can be collected in the price questionnaire. Prices for common health services can be obtained from the facility questionnaire. In the standard version of the health module, expenditures on treatment by provider type are gathered in Part E (Health Care Utilization and Expenditures). Part E in the expanded version of the health module gathers more precise measures, measures that can be used to construct prices by provider and by type and level of service, to calculate the costs of service in travel and waiting time, travel costs, and in-kind payments, and to adjust expenditures to account for insurance payments. Finally, information on travel time from Part F, which is included in both the standard and the expanded modules, is useful for calculating the time costs involved in obtaining treatment.

Question 19 in Table 8.1 regarding the impact of user fees involves estimating the effect of a change in prices on the use of health care services. A minimal analysis requires data on utilization and provider choice from Part E of the standard module, along with some measure of user fees that varies over the sample. Ideally the effects should be estimated from a fully specified health input demand equation, which requires data on utilization from Part E as well as data on all the variables that belong in the health input demand function. Because the questions essentially involve estimating price elasticities, particular attention should be paid to the construction of price variables.

User fees and the location of government health programs can have an effect on health outcomes (see Questions 20 and 22 in Table 8.1), which implies a need to collect data on health status. The difficulties involved in doing this were discussed above. As always, causal analysis requires estimation of a health demand equation, which in turn requires data on all the variables that belong in that equation.

Analyzing the impact of the location of government programs on the utilization of services, on households' choices of providers, on household expenditures on health care, and on government revenues (Question 21 in Table 8.1) requires, at minimum, data on the location of those programs. Of course, data on utilization rates and household expenditures are also needed to answer this question. Data on the location of programs are collected in the community questionnaire and in Part F of the household questionnaire (standard and

expanded versions). As discussed above, data on utilization and expenditures are collected in Part E of the standard and expanded versions of the health module for the household questionnaire.

Similar data are needed to analyze the impact of the location and availability of government services on health-related behavior (Question 23), using the health behavior data collected in Part B of the standard and expanded versions of the health module. Finally, the impact of alcohol and tobacco taxes on the consumption of those goods and on government tax revenues (Question 24) can proceed similarly. In particular, analysts could make use of data on taxes or prices that vary over the sample and data on consumption of the goods from Part B (Health-Related Behavior). Price elasticities can be obtained from estimates of demand functions for the goods, which again would require data on all of the variables in the demand equation.

**QUALITY OF HEALTH CARE SERVICES.** Policy issues concerning the quality of health care services are shown in Questions 25–27 in Table 8.1. There is really only one way to measure the quality of care, which is by using data collected in a health facility questionnaire. This information can be linked to the household using the information provided in Part F (Health Provider Knowledge) and the data on local health facilities in the community questionnaire. Sources for data on utilization rates, expenditures, and health status were discussed above.

**POLICIES REGARDING PRIVATELY PROVIDED HEALTH CARE SERVICES.** Questions 28–31 in Table 8.1 consider private sector health care providers and government policies toward them. The main issues are the quality of health care given by these providers and whether their services are in compliance with government regulations. Data from the facility questionnaire, in particular data collected from private health care providers, are essential (except for Question 30, which is discussed in the following paragraph). Parts A–F of the facility questionnaire collect data on quality of services as measured by the equipment and services available, while Parts G–K collect information on the process of care (that is, whether certain protocols are followed when examining patients). Data on prices, which are relevant for Question 29, are also collected in the facility questionnaire, and prices can also be

**Box 8.2 Cautionary Advice**

- *How much of the draft module is new and unproven?* The three versions (short, standard, and expanded) of the health module introduced in this chapter (and presented in Volume 3) are much larger than, and in most respects very different from, the health modules used in past LSMS surveys. But not everything is different; past LSMS surveys focused on utilization rates and household expenditures on health care, and since the short module and Part E of the standard and expanded versions of the module have the same focus, they are likely to work well because they are based on the experience of past surveys. However, most of the rest of the module collects data that have rarely been gathered in past LSMS surveys (such as data on self-reported health status, health-related behavior and insurance coverage). These parts are based mostly on the experience of non-LSMS surveys, but are also based to some extent on the 1990 Jamaica LSMS survey and the 1996 Brazil LSMS survey. These parts should work well but will probably require extra attention from survey designers. In particular, collecting data on self-reported health status, especially activities of daily living, is experimental. Health facility questionnaires have been administered in a variety of LSMS and non-LSMS surveys, and thus the design of the facility questionnaire is based on a large amount of experience.
- *How well has the module worked in the past?* The health module in past LSMS surveys worked well in achieving its main objective, the study of the choice of health care providers. Even so, the design was not well suited for cases in which individuals visited several health care providers during the reference period. The main criticism of past LSMS health modules concerns what they did not attempt to do. They collected almost no data on health status, health behavior, or insurance status. The standard and expanded versions of this module directly address these deficiencies.
- *Which parts of the module most need to be customized?* Several parts of the module need to be customized to reflect the circumstances in the country where the survey will be fielded. First, the types of facilities visited can vary widely across countries, so questions on the utilization of those facilities must be substantially modified. Second, the questions on health-related behavior must be adapted because of the sensitive nature of certain topics, as what is sensitive varies from country to country. Third, the activity list in the self-reported and the observed activities of daily living may need to be tailored to ensure that they are culturally appropriate. Fourth, several aspects of the facility questionnaire need to be changed to fit local conditions, since types of facilities, kinds of personnel, and services offered can vary substantially by country. After both the health facility questionnaire and the health module for the household questionnaire are customized, the draft questionnaires should be discussed in detail with officials from the ministry of health.

inferred from information on expenditures given in Part E of the standard and expanded versions of the health module for the household questionnaire.

The topic of Question 30 is whether private insurers are abiding by any government regulations that apply to them, such as whether they unlawfully drop individuals who have high medical costs or whether they comply with laws regarding their financial integrity. Panel data on insurance coverage may be able to detect whether household members' insurance coverage is illegally terminated, and to study similar issues regarding compliance with regulations pertaining to the treatment of insured individuals, but in general household survey data cannot provide information on the financial soundness of private insurance providers.

**HEALTH INSURANCE.** Questions 32–34 in Table 8.1 address policy issues related to health insurance. Information on who is covered by health insurance and what the specific benefits are (Questions 32 and

33) is provided in Part D of the standard and expanded health modules. A very small amount of information on the proportion of the sample population with insurance coverage is collected in the last two questions of the short module. Information on employer-provided health benefits is collected in the employment module, which is discussed in detail in Chapter 9.

Estimating the impact of insurance on utilization rates, expenditures, health behavior, health outcomes, and government revenues again requires estimation of the health demand function. As before, this requires data on all of the variables that belong in that function.

**HEALTH STATUS AND OTHER SOCIOECONOMIC OUTCOMES.** The last three questions in Table 8.1 (35–37) focus on how health status affects other socioeconomic outcomes. Answering each of these questions requires, at a minimum, data on health status and on the outcome of interest. The sources of data on health status were discussed above. A discussion of methods for estimating the determinants of consump-

tion, education and cognitive outcomes, agricultural productivity, and savings can be found in Chapters 5, 7, 19, and 20, respectively.

### **The Health Module**

This section introduces both a health facility questionnaire and three versions of a health module to be included in the household questionnaire. The three versions of the health module vary by length: short, standard, and expanded. Draft versions of all three health modules, as well as the draft facility questionnaire, are presented in Volume 3.

Table 8.2 summarizes the kinds of data collected in each version of the health module for the household questionnaire, and also shows the relevant data collected in the health facility questionnaire, the community questionnaire, the price questionnaire, and other modules of the household questionnaire.

#### *The Short Module*

The short version of the health module is designed for countries in which policymakers have some interest in health issues but health is not one of the most important policy topics in the survey. In this context, the short health module does four things. First, it collects some cursory information on the health status of household members. Second, it gathers information on all outpatient visits to health facilities by household members during the previous four weeks and on all inpatient visits in the previous 12 months. This information on utilization can be used to perform relatively simple analyses of the incidence of public health spending among different population groups.

Third, the short module collects some information on the household's expenditures on visits to the different providers. This provides a more accurate measure of the household's total expenditures on health than would be obtained by asking a single general question about these expenditures in the consumption module. This expenditure information also gives analysts a rough idea of the out-of-pocket cost of obtaining health care by type of service provider. (If more detailed expenditure data are required, the standard module should be used.) Fourth, the short module gathers very brief information on insurance coverage—just enough to indicate who is covered by insurance and who is not.

Even this short version of the health module may seem quite long. It contains 39 questions, about twice

as many as were used in the health modules of the first LSMS surveys in the mid to late 1980s. However, the health information gathered in those first LSMS surveys was of very limited use. Moreover, when the short module presented here is administered, most respondents will be asked far fewer than 39 questions, since most probably will not have visited a health care provider in the previous four weeks. The field test should verify that the time required to fill out the short health module in a typical household is 10–15 minutes.

A final point regarding the short module is that there is no need to administer a health facility questionnaire as well. However, it is still useful to collect some summary information about health facilities in the community questionnaire.

#### *The Standard Module*

The standard version of the health module should be used when health is one of the main policy issues in the country of the survey, but not necessarily the most important issue. The standard module collects data on self-reported health status (including self-reported activities of daily living), health-related behavior, child immunization, health insurance coverage, health care utilization and expenditures, and knowledge of local health care providers. Analysts can use these data to answer questions about the level and distribution of health policy outcomes, including how changes in policy affect various outcomes in the health sector. Of particular interest are estimates of the price elasticity of the demand for health care, which can be done using the standard or expanded version of the health module.

In most cases data on health service providers should be collected in the health facility questionnaire. Anytime the standard (or expanded) version of the health module is used, anthropometric data should also be collected, both for children and adults. The collection of anthropometric data is explained in detail in Chapter 10.

The average duration of an interview using the standard module should be approximately 5–10 minutes per individual. However, the length of these interviews is likely to vary substantially, since individuals who did not use any health care services during the previous 30 days will be able to skip many of the questions.

Finally, whenever the standard or the expanded health module is used, the survey team supervisor must

**Table 8.2 Health Data in the Health Module and in Other Modules**

Variable	Respondents	Short version	Standard version	Expanded version	Other module
<i>Adult: Self-reported</i>					
Self-reported health status	People 15 or older	1, 2	A3, A4, A11–A29	A3, A4, A11–A47	
Health-related behavior	People 15 or older		B1–B23	B1–B23	
Utilization and expenditures	People 15 or older	9–37	E1–E58	E1–E94	
Insurance coverage	People 15 or older	38, 39	D1–D17	D1–D17	
Employee health benefits	Employed individuals				Employment
<i>Adult: Objective measures</i>					
Body mass index	People 15 or older				Anthropometry
Directly observed activities of daily living	People 40 or older			G1–G11	
Cognitive functioning	People 40 or older			H1–H32	
<i>Child: Mother/guardian reported</i>					
Self-reported health status	People under 15	1–8	A3–A10	A3–A10	
Immunization	People under 6		C1–C11	C1–C11	
Utilization and expenditures	People under 15	9–37	E1–E58	E1–E94	
Insurance coverage	People under 15	38, 39	D1–D17	D1–D17	
<i>Child: Objective measures</i>					
Height/age and weight/height					Anthropometry
Birth weight					Fertility
Infant mortality					Fertility
Cognitive development					Education
<i>Household</i>					
Health-related behavior					Housing
Knowledge of health care providers	Wife of head		F1–F8	F1–F8	
<i>Other</i>					
Community health information					Community
Facility data	Director of facility				Facility questionnaire
Prices of medicines	Market vendors				Price questionnaire

Source: Authors' summary.

draw up a master list of codes for all of the health service providers in the community. This list is used to match different information from different parts of the questionnaire. In particular, these codes are entered in the section of the community questionnaire that gathers information on health facilities, in Part F of the standard module (which asks about the distance from the household to the nearest health facilities), on each health facility questionnaire, and in Part E (detailed utilization and expenditures) in the expanded module. Merging this information is very useful for policy analysis. For example, combining the information in Part F of the standard module with the data from the health facility questionnaire not only provides analysts with the distance between the household and the nearest health facilities but also yields a wide range of information about those facilities from the facility questionnaire.

### *The Expanded Module*

The expanded version of the health module takes the standard version as its starting point and adds more questions. It should be used in countries where health is the most important topic of the survey. Facility questionnaires should always be used when the expanded module is used. The expanded version goes beyond the standard version in two ways. First, it collects more data on health outcomes, specifically, observed data on mental health and data on activities of daily living and cognitive functioning. Second, it collects more detailed information on health care utilization and expenditures, which allows for more precise estimates of the demand for health and the demand for health care services. The detailed data gathered on utilization and expenditures includes information about the respondent's most recent visits to health care facilities, includ-

ing the purpose and cost of their visits, the treatment they received, their expenses on travel and medicines, and the time they spent traveling to the facility and waiting for treatment.

The amount of time required to complete the expanded version of the health module is approximately 10–20 minutes per individual. Again, this will vary widely across individuals, since some household members will have very few health problems and no recent visits to health facilities while others will have more of both.

#### *Hybrid Versions*

A final point regarding these three versions of the health module is that survey designers should feel free to put together “hybrid” versions that fit their specific data collection needs. For example, they could put together a module halfway between the short and standard versions or halfway between the standard and expanded versions. The standard and expanded versions are made up of several different submodules that are mostly independent of each other, so survey designers could construct a health module that consisted only of those submodules that covered issues of particular interest to policymakers in that country.

#### *Health Facility Questionnaire*

The draft health facility questionnaire is designed to collect information on the resources and practices of local health care facilities, both public and private. The information collected by this questionnaire can be used to construct health care prices and to assess the quality of care. Part A collects some basic information about the facility, such as its ownership, its sources of electricity and water, and its hours of operation. Part B asks about the medical services offered by the facility and the prices charged for them. Part C gathers information about health center employees. Part D collects data on medical equipment, including whether the equipment is in working order. In Part E the interviewer asks for a quick tour of the facility to evaluate the cleanliness and the availability of the equipment in the different rooms (for example, the examination room, the injection room, the vaccine storage area, and the room for doing laboratory tests). Part F asks whether different types of medicines are in stock. The information from Parts A–F can be used to construct measures of health care quality based on the structure of care.

In household surveys in which health issues are a top analytical priority, survey designers should also gather some facility data on the process by which health care is given. Such data can be gathered in Parts G–K. In each of these parts the interviewer describes a hypothetical patient with a particular illness or health care need. The respondent for the facility questionnaire is asked a series of questions to find out how such a patient would be treated by the health facility. The five scenarios presented to the respondent are:

- The treatment of an adult with a cough and a fever.
- The treatment of an infant with vomiting and diarrhea.
- A pregnancy examination.
- The provision of an IUD.
- The provision of oral contraceptives.

The precise scenarios used should vary by country depending on what health problems and issues prevail in each country. Developing new scenarios may be warranted in some countries, but survey designers doing so should consult a public health specialist who has experience with collecting this type of data.

In any country, health care service providers come in many shapes and sizes, from drug peddlers selling over-the-counter medicines to large urban hospitals. There are two types of approaches to designing questionnaires to fit the different kinds of providers. The first approach is to have one questionnaire that can be used for a wide variety of providers. The judicious use of skip codes can facilitate the use of such a questionnaire by instructing the interviewer not to ask questions that are irrelevant. The second approach is to design separate questionnaires for each kind of health care facility—a pharmacy questionnaire, a public clinic questionnaire, a private clinic questionnaire, and a hospital questionnaire. The draft facility questionnaire presented here is of the general type, but it can be modified and split into several different types if the second approach is used.

A final point regarding the facility questionnaire is how to select the facilities—a particularly important issue in large urban areas where there are usually many facilities from which to choose. This is discussed in the second section of this chapter and also in Chapter 13.

#### **Annotations to the Health Module**

This section contains specific notes on particular parts of the three versions of the draft health module intro-

duced in the previous section (and presented in Volume 3).

#### *Short Health Module*

**Q2.** This is a general measure of the health status of each individual. It is measured as the change over the past year to minimize bias caused by the influence of socioeconomic characteristics on respondents' assessments of their health status.

**Q3–Q8.** Diarrhea is most serious for children ages 0–6, so this is not asked of anyone 7 or older. If diarrhea is a serious problem for older people in a particular country, the age range can be changed accordingly. No questions are asked on any other specific diseases because, as explained in the second section of this chapter, subjective information on other diseases is likely to be very inaccurate and may well be biased.

**Q9–Q35.** These questions assume that there are six kinds of outpatient care providers—public hospitals, public health clinics, private hospitals or clinics, private doctors, private nurses/paramedics/midwives, and traditional health practitioners—and three kinds of inpatient care providers—public hospitals, public health clinics, and private hospitals or clinics. This can vary from country to country, so the number of both types should be adjusted depending on the circumstances in the country surveyed.

**Q11, Q14, Q17, Q20, Q23, Q26.** These questions exclude transportation costs (since transportation is not a direct form of medical care) but include purchases of medicines elsewhere. For some purposes survey designers may want to include transportation costs. In other cases policymakers may want more disaggregated expenditure information, such as separate questions on payments to the provider, payments for medicine purchased elsewhere, and payments for transportation. This would effectively add 12 more questions, and thus lengthen the interview time. This additional time may be worthwhile, depending on the interests of policymakers. Alternatively, even more detailed information could be gathered, along the lines of Part E in the standard module.

**Q21–Q26.** These questions assume that services from private individuals other than doctors and traditional

health practitioners are always outpatient services. If this is not the case, these questions must specify that information is only required for outpatient services, along the lines of Questions 9–11, which collect analogous information for outpatient visits to public hospitals. In addition, questions on inpatient services from private individuals and traditional health care practitioners should be added, along the lines of Questions 27–35 on inpatient services.

**Q38–Q39.** If insurance is relatively common, survey designers may want to collect more detailed information. In particular, it may be of interest to find out how much the insurer is reimbursing specific health care providers. (Note that household members are asked to report only costs that they pay, not costs paid by insurers.) This would entail asking additional questions for each of the six kinds of outpatient providers, as well as for the three kinds of inpatient providers. Again, whether this is worth the potential increase in interview time depends on the interests of policymakers. More detailed insurance questions are provided in Part D of the standard and expanded versions of the health module. One potential problem here is that a substantial proportion of respondents may not know how much the insurer pays if the costs are directly to the provider by the insurer, as opposed to reimbursing payments made by the individual. This can be checked during the field test of the draft questionnaire.

#### *Standard Health Module*

**A1–A2.** It is useful to know if the information is being provided by the person in question or if someone else is providing the information because the person in question is not available or is too young to answer for himself or herself. As discussed in Chapter 4, data collected directly from the person in question are likely to be much more reliable than data provided on this person's behalf by someone else.

**A4.** This is a general measure of the health status of each individual. It is measured as the change over the past year to minimize bias caused by the influence of socioeconomic characteristics on respondents' assessments of their health status.

**A5.** This question directs people to different questions depending on their age. Questions A6–A10 on diar-

rhea are usually only relevant for children ages 6 and younger. Further comments on Questions A6–A10 are provided in the previous subsection; see the comments on Questions 3–8 in the short module. Questions A11–A23 are only relevant for individuals ages 40 and older, and Questions A24–A29 are only relevant for individuals ages 15 and older.

**A11–A29.** These questions address the self-reported activities of daily living. The particular self-reported activities of daily living that should be collected, and how they should be interpreted, will vary across countries. The easiest activities of daily living are those in Questions A11–A16 (dressing oneself, standing from a sitting position, using a toilet). As indicated in Question A17, anyone who has difficulty doing any of these activities need not answer the remaining questions because he or she will not be able to do these more strenuous activities.

**PART B.** This part contains questions on smoking, alcohol consumption, and exercise, as well as on knowledge of the health consequences of sexual behavior. The particular behavior to be examined and the specific way in which this behavior manifests itself will vary over countries; for example, in countries where travel by car or motorcycle is common, survey designers may want to ask questions about use of seatbelts and motorcycle helmets. Because of this, Part B may need to be substantially modified to reflect the circumstances that prevail in certain countries. Ideally it would be useful to ask questions not only on knowledge of the consequences of sexual behavior but also on respondents' sexual behavior, particularly risky sexual activities. However, in most countries the sensitive nature of these activities makes it difficult to ask such questions. See the second section of this chapter for further discussion.

**B2–B13.** In addition to the kinds of tobacco asked about here, survey designers might choose to ask about the use of pipe tobacco, if pipe smoking is common in the country surveyed.

**B7, B8, B13, B15, B16.** In addition to asking about quantities consumed, survey designers may want to ask about expenditures. Such questions should be asked immediately after the questions on quantities. This may provide more accurate information on expenditures than would be obtained in the consumption

module. Questions about expenditures are not included in this version because price data are collected in the price questionnaire for tobacco and alcohol; for most purposes analysts could multiply those prices by the quantities provided here to obtain a reasonably accurate measure of expenditures on alcohol and tobacco. The price information on tobacco and alcohol collected in the price questionnaire should be in the same units used here for quantities consumed.

**B15–B16.** If bottles vary in size, it may be better to ask this question in terms of ounces.

**PART C.** These questions ask about the immunization status of all children younger than 7. As explained in Chapter 3, this information could also be collected in the anthropometry module or the fertility module. If the standard version of the health module is used, it is more convenient to collect this information here than in the anthropometry module, since anthropometry measurements are often done by individuals who are not trained interviewers. If there is a fertility module in the survey, it is useful to collect immunization data in the fertility module because such data will also be collected for children who have died.

**C1–C2.** Mothers are asked questions about their children's vaccination history based on their vaccination cards and are asked to recall whether any information does not appear on the card. If the card is unavailable or incomplete, the mothers are asked to recall their children's entire vaccination history.

**C4, C6, C8.** The way these immunizations are given will vary by country. Survey designers should consult knowledgeable officials in the ministry of health to find out how immunizations are given, and change the wording of these questions accordingly.

**PART D.** This schedule asks about any insurance coverage that the members of the family may have, the source of the coverage, and the benefits provided.

**D1.** Health insurance includes health care provided by the government or an employer without charge or for a reduced charge. Some respondents may not think of this as insurance. In this case the wording of Question D1 must clearly indicate that such health care is considered to be insurance.

**D3.** In some countries it is possible for household members to have more than one kind of insurance. For example, insurance in the form of free health care provided to government workers may be deemed inadequate, and in response some of them may also visit private health care providers and purchase private health insurance. In this case Part D must be expanded to allow for two kinds of insurance per household member, which in general will entail repeating Questions D5–D17 for a second type of insurance.

**PART E.** This part collects basic information on self-treatment and on inpatient and outpatient treatment by type of facility during the previous four weeks, as well as on the total costs associated with each type of treatment. These questions assume that there are six kinds of outpatient care providers—public hospitals, public health clinics, private hospitals or clinics, private doctors, private nurses/paramedics/midwives, and traditional health practitioners—and three kinds of inpatient care providers—public hospitals, public health clinics, and private hospitals or clinics. This may vary from country to country, so the number of both types should be adjusted depending on the circumstances of a given country.

**E5, E11, E17, E23, E29, E35, E41, E47, E53.** One easy way to check whether a person is insured is to add another column onto the fold-out sheet with the names of all household members (see Chapters 3 and 6) that indicates whether the person is covered by insurance.

**E25, E26, E31, E32.** These questions assume that services from private individuals and traditional health practitioners are always outpatient services. If this is not the case, these questions must specify that information is only required for outpatient services along the lines of Questions E1 and E2, which collect analogous information for visits to public hospitals. In addition, questions will have to be added that ask for inpatient visits (hospitalizations) with these kinds of health care providers.

**PART F.** These questions ask about providers that are known to the respondent. The respondent should be the adult woman in the household who knows the most about local health care providers. In most cases this will be the wife of the head. If no adult women

live in the household or if an adult male is clearly more knowledgeable than any of the adult women, the most knowledgeable adult male should be interviewed. This information can be used to measure access to health care and to estimate health care prices.

The kinds of health facilities from which to collect this information will vary from country to country. For the health facilities that are most common, the nearest two or three (within some radius of the center of the community) should be covered. The questionnaire lists only three facility types—public hospital, public clinic, and private hospital or clinic—but in most cases more should be included, such as private pharmacies, itinerant drug peddlers, traditional healers, and family planning centers. The specific types to list should be discussed with officials from the ministry of health.

**F2–F3.** The codes associated with the names of the facilities should be taken from a master list of health facilities in the community. This list is drawn up by the survey team supervisor, as explained in the third section of this chapter.

#### *Additional Questions for the Expanded Health Module*

**A30–A47.** The wording of these questions, which pertain to mental health, will have to be adapted to reflect the prevailing culture in the country of the survey. In some countries the prevailing culture may prevent people from answering these questions accurately, in which case there is little use in asking them.

**PART E.** This version of Part E is nearly twice as long as the version in the standard module because it collects more detailed expenditure and utilization data. This level of detail allows for a wide variety of descriptive analyses, and can also be used to construct prices in the health and health input demand equations. The schedule collects information on the particular provider used by the respondent, along with the purpose of the respondent's visit and measures of time costs, distance, types of service received, and monetary or in-kind payments made for the service. The types of provider will vary from country to country.

**E1–E57.** This long set of questions can accommodate up to three visits in the previous four weeks. If, say, more than 1 percent of the sample have had more than three

visits, questions could be asked about up to four visits. In either case there will inevitably be a few individuals whose visits exceed the number that the questionnaire can accommodate. For these respondents, extra blank sheets or a blank questionnaire should be used to record all the visits. The data entry program should be designed to allow for a large number of visits.

**E58–E90.** These questions cover inpatient stays (hospitalizations) in the previous 12 months, allowing up to 3 inpatient stays per individual. See the comment above regarding Questions E1–E57 for what to do when individuals make more than three inpatient stays.

**PART G.** Interviewers should ask all people 40 and older to perform directly-observed activities of daily living as one measure of their health status. For more detail on the protocols used see Guralnik and others (1989).

**G3–G5.** These questions measure difficulty in standing from a sitting position. Individuals who can rise from a chair without using their arms answer the first part of Question G3 and proceed to G4, where they are timed on their ability to do this five times. They then proceed to G6. Individuals who are unable to rise from a chair without using their arms are asked to try doing so using their arms. If they can do so, they proceed to Question G5, in which they are asked to do this as many times as they can, up to five times. If they cannot do so they proceed to G6.

**G9.** In this activity participants are asked to pick up a pencil from the floor and return to a standing position. They may bend over or squat or do whatever they like to pick up the pencil.

**G10.** The respondent is asked to tap his or her foot 10 times for each foot. For those who can tap their foot 10 times, the time is recorded (and the number of taps is recorded as 10). For those who cannot do this 10 times, the number of times they can do it is recorded, and no time is recorded.

**G11.** The respondent is asked to hold his or her arm straight out in front and then slowly raise it up over his or her head. He or she should continue the motion, keeping the arm straight while rotating the shoulder as

needed, until the arm is extended behind with the palm down. If the person has to bend his or her arm to do this, the “partial” response should be coded.

**PART H.** These questions on cognitive functioning should be administered only to people aged 40 and older. They should be administered by the interviewer. They are based on tests used in the Bangladesh Family Life Survey (which was supported by the Rand Corporation). Some of the questions are culturally specific, so pretesting will be needed. Some respondents may find the questions so simple that they are annoyed or insulted, yet most such respondents should be satisfied by an explanation of the purpose of the questions.

**H8.** For countries in which there is no prime minister or the prime minister is not the best-known political leader, replace “prime minister” with “president” or a similar appropriate term.

**H9.** In most countries “orange” and “cat” should be appropriate, but in some countries another common fruit and animal may need to be used. “House” should be appropriate in virtually any country.

**H10, H11, H14, H15.** The term “dollars” should be replaced with the name of the local currency. In Question H14, “dime” should be replaced with the appropriate analogous term.

**H11–H13.** In countries in which rice is not a common staple, the word “rice” should be changed to a common staple grain or tuber.

#### *Facility Questionnaire*

As explained in the third section of this chapter, Parts A–F of the facility questionnaire collect information on the structure of health care, while Parts G–K collect information on the process of health care. The first type of information is essential, but the second type is more experimental and thus may be regarded as optional. If resources are scarce, Parts G–K can be dropped, but if financial constraints are less tight, it will be useful to collect this information.

**COVER PAGE.** The code number for the health facility can be obtained from the master list of local health facilities, which in most cases will be collect-

ed in the community questionnaire. See the discussion of sampling issues in the second section of this chapter.

**A4.** It is important for analysts to know when a facility first opened to understand how its appearance in the community may have changed health outcomes. As discussed in Chapter 23, methods of analysis that use panel data often use this information.

**A13–A14.** Registration fees are general fees charged to each individual who visits a clinic, regardless of what treatment he or she receives. In most cases extra fees are added for particular services. In countries in which registration fees do not exist, these two questions can be dropped.

**PART B.** The list of the types of services offered must be modified to reflect the services in the country of the survey. For example, in East Asian countries survey designers may need to add acupuncture.

**PART C.** If the questionnaire is applied to a large facility such as an urban hospital, it will not be practical to ask questions about each individual. In this case a set of questions should be asked about how many employees or different types of employees (for example, doctors, nurses, or technicians) work there. To get more detailed information on hours worked, a random sample of perhaps four doctors, four nurses, and four technicians could be drawn up, and all of the questions in Part C could be administered to them.

**C2, C5.** The types of medical personnel and the kinds of degrees that they may have must be modified to fit the circumstances of the country. Survey designers should consult officials at the ministry of health when drawing up these lists.

**C11–C12.** These questions are intended to find out the extent to which health workers in a public facility also provide services in some kind of private practice. This is an important but sensitive issue, so these questions need to be field tested and used with care. It may not be necessary to ask these questions for private facilities, but in some countries even workers in private facilities may have separate private practices.

**PART D.** The types of equipment listed here must be modified to fit the country in which the survey is being carried out.

**PART E.** The text in lowercase consists of the statements and questions made by the interviewer to the person being interviewed. The text in uppercase is not to be read aloud. Instead, the interviewer needs to fill in this information based on what he or she observes. As much as possible, the interviewer should not indicate to the person being interviewed what he or she is observing.

**E1, E2, E4, E20, E21.** During the training of the interviewers, clear definitions need to be established regarding what constitutes “clean” and what constitutes “dirty.” This should be included in the interviewer manuals prepared for the team members who fill out this form.

**PART F.** The list of medicines must be adjusted to fit the ones commonly used in the country. Survey designers should consult officials at the ministry of health when drawing up this list.

**PARTS G–K.** The scenarios used must be modified to fit the particular characteristics of health and health care in the country. This can be done only by an expert in public health, preferably one with experience gathering this type of information.

### **Appendix 8.1 Estimating the Effects of Policy on Outcomes**

Many of the policy questions discussed in this chapter involve estimating the effect of a change in some health policy, such as the price or quality of health care, or taxes on health related goods, on policy outcomes such as utilization of (or expenditures on) health care services, health-related behavior—or on health itself. Answering questions of the form “what is the effect of a change in some variable on an outcome?” involves undertaking a causal analysis. Specifying data requirements for this type of analysis is not simple. There are often a variety of options for choosing the set of variables required for the analysis, and evaluating the merits of each of these options requires an understanding of the methodological issues involved. This appendix is designed to provide detail

that goes beyond the brief methodological discussion in the second section of the chapter. Readers interested in collecting data for the purpose of undertaking a causal analysis should also consult Chapter 26 on economic models and econometric methods and Chapter 23 on panel data.

The approach for estimating the effects of a change in a policy on a health outcome involves multiple regression analysis in which the dependent variable is some measure of utilization of health care, demand for health-related goods, health-related behavior, or (change in) health status. An economic model of the demand for health provides guidance for determining which variables are to be included as explanatory variables in the regression equation.

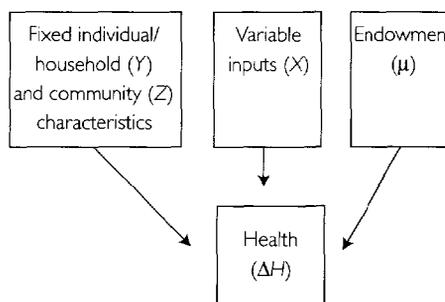
The foundation of this model is the health production function, which describes the biomedical process through which inputs (for example, medical care, nutrition, and smoking) are transformed into health. The health production function is illustrated in Figure A8.1; examples of inputs for adult, child, and newborn health production functions are listed in Table A8.1. The inputs into the health production function are only those factors that directly affect health, such as nutrition and medical care. Factors that indirectly affect health by altering behavior, such as medical care prices, are not inputs.

Following Grossman (1972), health can be thought of as a form of human capital. An individual's health stock at any point in time is determined by an initial genetic endowment, subsequent behavioral choices (for example, diet, medical care, smoking, and exercise), and factors that are beyond the control of the household.

Over a given period of time the change in an individual's health status is determined by the health production function. The health production function makes explicit the mechanism that transforms inputs consumed during a period of time into changes in health during that period of time. There are several types of factors that directly influence health:

- Health input choices such as food (diet), medical care, and physical activities.
- Individual behaviors that are not motivated by health considerations but nevertheless affect health, such as smoking and engaging in risky sexual activity.
- Household behavior that affects health by affecting the household disease environment, including cooking, sanitation and waste disposal practices, and water supply decisions.

**Figure A8.1 The Health Production Function**



Source: Authors' summary.

- Fixed individual and household characteristics that affect the productivity of health inputs, such as education, age, gender and family background. For example, better-educated individuals may be better able to follow a recommended medical treatment and thus obtain greater improvements in health from medical services.
- Fixed community characteristics that are exogenous to (beyond the control of) households, such as quality of available medical services and community programs concerning sanitation, water, and vector control.

The first three types of factors are variable health inputs. The fourth type (individual and household characteristics) and the fifth type (community characteristics) include fixed inputs and all other inputs into the production function that the household cannot control. The most important community variables are government policies that directly influence health outcomes (and thus are inputs into the production function) and policies that indirectly affect health outcomes by influencing other input decisions. For example, environmental efforts to reduce pollution have a direct influence on health and belong in the health production function. But these policies also have an indirect influence on health because pollution affects the risk of getting ill and consequently the utilization of preventative care. Similarly, the quality of care affects health outcomes both directly, by changing the marginal product of medical care, and indirectly, by inducing individuals to seek additional care.

The health production function does not include all policies that affect health outcomes. A number of policies have no direct impact on health outcomes but have indirect effects through their influence on the

**Table A8.1 Inputs and Other Variables in Health Production Functions**

Input	Adult	Child	Newborn
<i>Variable inputs</i>			
Nutrition	Food	Food	Mother's food during pregnancy
Physical activity	Amount and intensity of work	Amount and intensity of work	Mother's amount and intensity of work during pregnancy
<i>Health care inputs</i>			
	Preventative checkups	Preventative checkups	Prenatal care
	Immunization	Immunization	Delivery
	Inpatient curative care	Inpatient curative care	
	Outpatient curative care, drugs	Outpatient curative care, drugs	
<i>Individual health-related behavior</i>			
	Smoking	Smoking	Smoking of mother
	Alcohol consumption	Alcohol consumption	Alcohol consumption of mother
	Safe sex practices	Safe sex practices	Safe sex practices of mother
<i>Household health-related behavior</i>			
	Sanitation practices	Sanitation practices	Sanitation practices
	Cooking practices	Cooking practices	Cooking practices
	Waste disposal practices	Waste disposal practices	Waste disposal practices
	Water supply decisions	Water supply decisions	Water supply decisions
<i>Fixed inputs and characteristics</i>			
<i>Individual- and household-level</i>			
	Age	Mother's age	Mother's age
	Education	Mother's education	Mother's education
	Gender	Gender	Family background
	Family background	Family background	
<i>Community-level</i>			
	Sanitation	Sanitation	Sanitation
	Water	Water	Water
	Vector control	Vector control	Vector control
	Public education	Public education	Public education
	Quality of care	Quality of care	Quality of care

Source: Authors' summary.

choice of variable inputs. Most of these purely indirect policies are financial variables that affect the cost of the inputs. For example, the monetary price (including insurance coverage benefits), travel costs, and opportunity cost of time affect how much preventative and curative medical care an individual can obtain. Another indirect policy is information about prices and the availability of care. (Information about the productivity of care may directly enter the production function as it influences the efficiency of the input.) Other policies typically have both direct and indirect effects. The distinction between policies with direct and indirect effects has important implications for the methods used to measure the impact of policy measures on health outcomes.

The health production function can be expressed in terms of an equation. The amount of health produced during a given time period is:

$$(A-1) \quad \Delta H = H(X; Y, Z, \mu)$$

where  $\Delta H$  is the change in health over the time period,  $X$  is the set of variable inputs chosen during that period,  $Y$  is the set of individual/household characteristics beyond the control of the household,  $Z$  is the set

of community characteristics that directly affect health, and  $\mu$  is the individual's health endowment at the beginning of the period.

In the health production function, the dependent variable,  $\Delta H$ , is the change in health status over the period. The health production function shows how behavior during the period ( $X$ ), exposure to community environment ( $Z$ ), and factors beyond the control of the household ( $Y$ ) affect the flow of (that is, the changes in) health. If the left-hand side were the stock of health,  $H$ , all of the input decisions and exposure to environment factors over the individual's lifetime would belong on the right-hand side. This has important implications for the measurement of health status because most of the measures are really stocks of health as opposed to flows.

In principle, it is possible to apply multiple regression analysis to data on health and on all of the inputs and individual and community characteristics to estimate the effects of each of the inputs on health. That is, it is possible to estimate the health production function. While estimates of the health production function can be useful in improving our understanding of health production technology, direct estimation of the production function is not the ideal way to evaluate

the effectiveness of policies, for two reasons. First, to obtain reliable estimates of the effects of inputs on outputs it is usually necessary to have data on all of the inputs, which is rarely possible in practice. Second, the production function approach captures only direct effects; it does not allow for estimates of the effects of indirect policies, such as policies that affect prices, on health.

An alternative is to estimate reduced form health demand and health input demand equations. The reduced form approach (illustrated in Figure A8.2) is more useful for policy evaluation purposes because it captures both the direct and indirect effects of policies on outcomes. The reduced form approach models the effect of all fixed factors on both input choices and health outcomes.

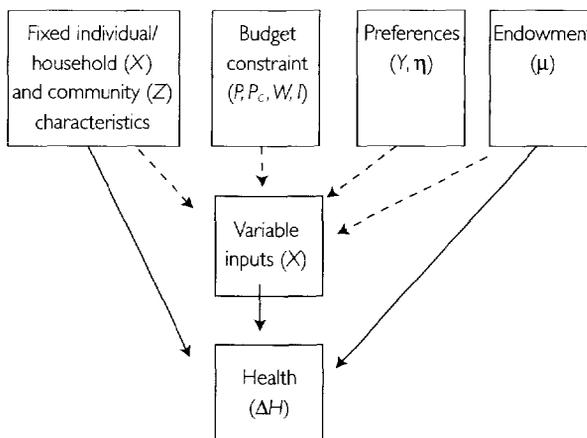
Consider first the household's choice of variable inputs, which depend on the household's preferences for health relative to other goods and leisure, the household's budget constraint, other fixed inputs and the endowment.<sup>9</sup> Preferences are related to observable household characteristics, such as education, as well as to unobservable characteristics. The budget constraint is determined by the household's income ( $I$ ) and all of the prices that it faces, including the prices of goods that directly affect the health production function ( $P$ ), the prices of nonhealth consumption goods ( $P_C$ ), and the market wage rate ( $W$ ). Indirect policies and prices affect input choices through the budget constraint. So which variable health inputs to use, and how much of each input to use, depends on all of the exogenous variables potentially observable to researchers—household characteristics, prices, wage rates, income, and fixed household- and community-level inputs—and two factors that cannot be directly observed by researchers: the household's preferences ( $\eta$ ) and the individual's health endowment ( $\mu$ ). The relationship between these observable exogenous variables and the demand for a variable health input is the health input demand equation for that input.<sup>10</sup>

The health input demand equation is described by the dashed arrows in Figure A8.2 and, in equation form, is:

$$(A-2) \quad X_n = X_n(P, P_C, W, I, Y, Z, \mu, \eta)$$

where  $P$  is the set of prices for the health inputs,  $P_C$  is the price of consumption goods not related to health,  $W$  is the wage rate,  $I$  is income,  $Y$  is the set of fixed

Figure A8.2 The Health Input Demand Function



Source: Authors' summary.

individual and household level characteristics,  $Z$  is the set of community characteristics,  $\mu$  is the health endowment from the production function, and  $\eta$  reflects unmeasured preferences in favor of allocating resources to health. Both direct and indirect policies are included in the determinants of input demands.

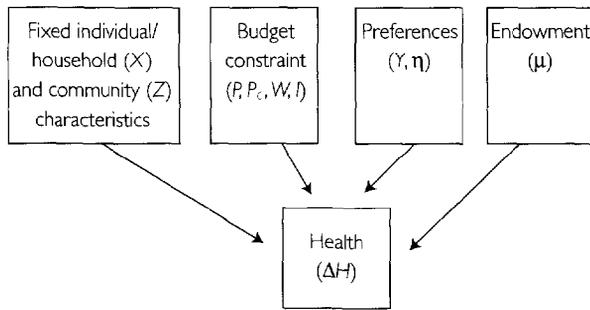
It is from the input demand equations that we are able to determine the effects of policies such as public health and education programs and prices (all of which are community characteristics) on health care utilization and health-related behavior. Because prices are defined as the effective price paid by the consumer, they incorporate taxes, subsidies, travel expenses, and insurance copayments as well as the prices charged by health facilities. The effects of travel and waiting times can also be included in the effective price by allowing time costs to enter the calculation of the price. Alternatively, distance can be specified as a distinct type of price variable. Estimates of the effects of policy changes on expenditures are obtained by computing the effects of each change in policy on utilization by type of service, and multiplying the level of utilization of each service by the price of each service.

Using a health demand function (Figure A8.3) it is also possible to make a direct estimate of the effects of the observed exogenous variables, including both direct and indirect policies, on health outcomes.

The health demand function is obtained by substituting the input demand functions in equation A-2 into the health production function:

$$(A-3) \quad \Delta H = HD(P, P_C, W, I, Y, Z, \mu, \eta).$$

**Figure A8.3 The Health Demand Function**



Source: Authors' summary.

Both direct and indirect policies are included in the health demand function whereas only the direct policies were included in the health production function. Estimates of the health demand equation are used to estimate the effects of both the direct and indirect policies on health status.

Under ideal circumstances, answers to most questions regarding the effect of policies on outcomes can be obtained through multiple regression analysis, based on equations A-2 and A-3, with cross-section data. However, there are several different estimation problems that complicate the empirical work. These methodological problems can often be addressed using techniques that require panel data (data from repeated surveys that interview the same households) or other additional variables (such as instrumental variables). For further discussion see Chapters 23 and 26.

**Dynamic Issues**

In the static model presented above, it was assumed that the household's choices are made at one point in time and that the stock of health in the previous period is exogenous and represented by  $\mu$ . An alternative to this is a dynamic formulation of the health production function in which the change in health over a period of time is determined in part by the health at the beginning of that period. In other words, the efficacy of medical care depends in part on the severity of sickness and general frailty. The dynamic health production function can be expressed as:

$$(A-4) \quad H_t - H_{t-1} = H(H_{t-1}, X_t, Y_t, Z_t, \epsilon_t).$$

Unmeasurable random factors (shocks), varying over time, that affect the change in health are reflect-

ed in  $\epsilon_t$ . The dynamic health production function is recursive in health since the lagged stock of health is a determinant. That is, all of the previous years' input choices plus the lifetime of exposures to environmental risks and the individual's genetic health endowment are subsumed in  $H_{t-1}$ . Repeated substitution of the health production function into equation A-4 for  $H_{t-1}$ ,  $H_{t-2}$ ,  $H_{t-3}$ , and so on, until lagged health variables are no longer on the right-hand side, yields an equation in which the change in health during this time period is a function of lifetime input choices, individual/household and community characteristics, the shocks, and the endowment of health:

$$(A-5) \quad H_t - H_{t-1} = H(X_t, \dots, X_0; Y_t, \dots, Y_0; Z_t, \dots, Z_0; \epsilon_t, \dots, \epsilon_0).$$

If the dynamic production function is appropriate, the input demand and health demand equations should also contain lagged health, as well as lagged household wealth,  $A_{t-1}$  (which reflects a potential source of income). The input demand equations, then, are:

$$(A-6) \quad X_n = X_n(H_{t-1}, A_{t-1}, P, P_c, W, I, Y, Z, \mu, \eta, \epsilon)$$

and the health demand equation is:

$$(A-7) \quad HD(H_{t-1}, A_{t-1}, P, P_c, W, I, Y, Z, \mu, \eta, \epsilon).$$

This approach makes it possible to see how an individual's health status modifies the impact that other determinants of health have on health status. More importantly, by conditioning on lagged health status, the reduced form health and input demand functions depend only on contemporaneous resources and policies, as opposed to past resources and policies, because the rest of the individual's health history is captured by  $H_{t-1}$ .

If it is not possible to condition on lagged health status, the interpretation of the model is quite different and indeed there may be serious omitted variable bias. To obtain a specification of equations A-6 and A-7 that does not depend on lagged health, all of the determinants of lagged health must be substituted into equations A-6 and A-7. These include the whole history of resources and policies. Therefore, the model should include not only current values of  $P, P_c, W, I, Y,$  and  $Z$ , but also the whole history of these variables.

This discussion of dynamics highlights the problem of modeling stocks versus modeling flows. An individual's stock of health is a function of a lifetime of decisions and exposure to environmental factors. Therefore, in the reduced form, health depends on the influences of the entire history of the exogenous variables. The flow or change in health from one period to the next depends on current influences and health status at the beginning of the period. Estimation of the health and health input demand equations in this framework poses the additional requirement that lagged health and wealth data be collected. This is only possible with true panel data.

## Notes

The authors gratefully acknowledge financial support from the US National Institute of Aging and comments from Girindre Beeharry, Ricardo Bitran, Eduard Bos, Mariam Claeson, Margaret Grosh, Gerry Hendershot, Peter Heywood, Prabhjat Jha, Edna Jones, John Peabody, Alexander Preker, and Laura Shrestha. Research assistance was provided by Debra Fogarty.

1. Primary health care is the provision of basic outpatient services, such as services one would find in a simple health clinic. Secondary health care provides more advanced outpatient services and also provides several kinds of inpatient services. This level of health care would be offered by large clinics and regional hospitals. Tertiary health care pertains to the best hospitals in the country, which may provide specialized services and may serve as teaching hospitals (hospitals in which doctors are trained and medical research is conducted).

2. The price elasticity is the percentage change in an outcome variable (in this case, use of health care services) that results from a given percentage change in the price of that variable.

3. The specific data needed would be the total cost plus the costs of each of the following: wages (and other staff benefits), drugs, supplies, maintenance, utilities, and capital costs. The facility questionnaire introduced in the third section of this chapter (and presented in Volume 3) does not collect such data because in many countries budgeting is done at a higher level, so that heads of facilities would not be able to answer these questions. In such cases the data can usually be collected from the ministry of health at a district or regional level.

4. Employer-provided health benefits in the form of subsidized health care are essentially a form of insurance. Thus the discussion of insurance applies to those kinds of benefits.

5. One could try to collect the information needed to analyze insurance issues on the small number of people who have had catastrophic illnesses by using calendars (forms on which the respon-

dent fills in utilization data for each day) or conducting a special survey for the subsample of respondents who report severe illness or an illness that started before or ended after the recall period. The issues that arise in developing such surveys are similar to those that arise in other studies involving calendars, such as time use and food consumption studies. Such questionnaires are beyond the scope of this chapter.

6. There are other problems as well. If the number of services or number of health care providers is large but the number of households in the community that are interviewed is small, one may not get information for some kinds of services or for some providers because none of the sampled households are familiar with them. This is particularly the case if one uses information only from households that actually used a service. Another problem with collecting data only from households that used a service is that the health care facilities covered are not necessarily a random sample of the population, which could lead to biased estimates of the characteristics of health service providers and community health programs.

7. Many households may not have very accurate knowledge of the distances to local facilities. One alternative is to use global positioning system devices to pinpoint the latitude and longitude of each household and each health facility.

8. A third relationship is a health production function. However, this is more difficult to estimate and does not account for indirect effects of health policies. Thus, for most purposes, such estimates are not recommended. For further discussion see Appendix 8.1.

9. Formally, the household maximizes utility, which is a function of health, leisure, and consumption of other goods subject to the budget constraint and the health production function.

10. Technically this would be a conditional health input demand equation if some of the inputs were chosen by the household, which may be the case for household income. For the same reason, equation A-3 may be a conditional health demand equation.

## References

- Alba, Michael. 1998. "A Hospital Cost Function for the Philippines." In Alex Hennis, ed., *Health Sector Reform in Asia*. Manila: Asian Development Bank.
- Andrews, G., A. Esterman, A. Braunack-Mayer, and C. Rungie. 1986. *Aging in the Western Pacific*. Manila: World Health Organization.
- Behrman, Jere, and Victor Lavy. 1992. "Child Health and Schooling Achievement: Causality of Association." University of Pennsylvania, Department of Economics, Philadelphia, Penn.
- Bound J. 1991. "Self-Reported versus Objective Measures of Health in Retirement Models." *Journal of Human Resources* 26 (1): 106–38.

- Bound J., M. Schoenbaum, and T. Waidman. 1995. "Race and Education Differences in Disability Status and Labor Force Participation." *Journal of Human Resources* 30 (supplement): S227–S267.
- Brook, R.H., and others. 1983. "Does Free Care Improve Adult Health? Results from a Randomized Controlled Trial." *New England Journal of Medicine* 309 (23).
- Deaton, Angus. 1988. "Quantity, Quality and the Spatial Variation in Price." *American Economic Review* 78 (3): 418–30.
- de Ferranti, David. 1985. "Paying for Health Services in Developing Countries: An Overview." Staff Working Paper 721. World Bank, Washington D.C.
- Deolalikar, A. 1988. "Nutrition and Labor Productivity in Agriculture: Estimates for Rural South India." *Review of Economics and Statistics* 70 (3): 899–32.
- Dow, William. 1996. *Unconditional Demand for Health Care in Côte d'Ivoire: Does Selection on Health Status Matter?* Living Standards Measurement Study Working Paper 127. Washington, D.C.: World Bank.
- Dow, William, Paul Gertler, Robert Schoeni, and Duncan Thomas. 1997. "Health Care Prices, Health, and Labor Outcomes: Experimental Evidence." RAND Corporation, Santa Monica, Cal.
- Ferraro, Kenneth, and Melissa Farmer. 1999. "Utility of Health Data from Social Surveys: Is There a Gold Standard for Measuring Morbidity?" *American Sociological Review* 64 (2): 303–15.
- Gertler, Paul, and Jonathan Gruber. 1997. "Insuring Consumption against Illness." NBER Working Paper 6035. National Bureau of Economic Research, Cambridge, Mass.
- Gertler, Paul, and Jeffrey Hammer. 1997. "Strategies for Pricing Publicly Delivered Health Care Services." In G. Schrieber, ed., *Innovations in Health Care Financing*. World Bank: Washington D.C.
- Gertler, Paul, and Roland Sturm. 1997. "Private Health Insurance and Public Expenditures in Jamaica." *Journal of Econometrics* 77 (1): 237–57.
- Gertler, Paul, and Jacques van der Gaag. 1990. *The Willingness to Pay for Medical Care: Evidence from Two Developing Countries*. Baltimore, Md.: Johns Hopkins University Press.
- Gertler Paul, and J. Zeitlin. 1996. "The Returns to Childhood Investments in Terms of Health Later in Life." RAND Corporation, Santa Monica, Cal.
- Gilleskie, Donna. 1998. "A Dynamic Stochastic Model of Medical Care Use and Work Absence." *Econometrica* 66 (1): 1–45.
- Glewwe, Paul, and Hanan Jacoby. 1995. "An Economic Analysis of Delayed Primary School Enrollment in a Low Income Country: The Role of Early Childhood Nutrition." *Review of Economics and Statistics* 77 (1): 156–59.
- Glewwe, Paul, Hanan Jacoby, and Elizabeth King. Forthcoming. "Early Childhood Nutrition and Academic Achievement: A Longitudinal Analysis." *Journal of Public Economics*.
- Goldman, Fred, and Michael Grossman. 1978. "The Demand for Pediatric Care: An Hedonic Approach." *Journal of Political Economy* 86 (2): 259–80.
- Graham, W, W. Brass, and R. Snow. 1989. "Estimating Maternal Mortality: The Sisterhood Method." *Studies in Family Planning* 20 (3).
- Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy* 80 (2): 223–55.
- Guralnik, J., L. Branch, S. Cummings, and J. Curb. 1989. "Physical Performance Measures in Aging Research." *Journal of Gerontology* 44: 141–46.
- Hammer, Jeffrey. 1997. "Economic Analysis for Health Projects." *World Bank Research Observer* 12 (1): 47–71.
- Idler, Ellen, and Yael Benyamini. 1997. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior* 38 (1): 21–37.
- Ju, A., and G. Jones. 1989. *Aging in ASEAN and its Socioeconomic Consequences*. Singapore: Institute of Southeast Asian Studies.
- Kington R., and J. Smith. 1997. "Socioeconomic Correlates of Adult Health." *Demography* 34 (1): 159–170.
- Kochar, A. 1995. "Explaining Household Vulnerability to Income Shocks." *American Economic Review* 85 (2): 159–64.
- . 1997. "Explaining Poverty: An Empirical Analysis of the Effects of Ill-Health and Uncertainty on the Savings of Rural Pakistani Households." Stanford University, Department of Economics, Stanford, Cal.
- Lewis, Maureen, and A. Medici. 1995. "Private Payers of Health Care in Brazil: Characteristics, Costs, and Coverage." *Health Policy and Planning* 10 (4).
- Litvack, Jennie, and Claude Bodart. 1993. "User Fees Plus Quality Equals Improved Access to Health Care: Results from a Field Experiment in Cameroon." *Social Science and Medicine* 37 (3).
- Macro International. 1995a. "Model "A" Questionnaire: With Commentary for High Contraceptive Prevalence Countries." Demographic and Health Surveys, Calverton, Md.
- . 1995b. "Tanzania: Knowledge, Attitudes and Practices, 1994." Demographic and Health Surveys in collaboration with the Bureau of Statistics in Tanzania, Calverton, Md.
- McDowell, Ian, and Clair Newell. 1996. *Measuring Health: A Guide to Rating Scales and Questionnaires*. Oxford: Oxford University Press.
- Peabody, John. 1996. *Quality of Care: Key Policy Issues*. Vietnam National Health Survey Background Paper. RAND Corporation, Santa Monica, Cal.
- Peabody, John, Paul Gertler, and Arleen Leibowitz. 1998. "The Effects of Structure and Process of Medical Care on Birth Outcomes in Jamaica." *Health Policy* 43 (1): 1–13.

- Peabody, John, Jeff Luck, Peter Glassman, Timothy Dresselhaus, and Martin Lee. Forthcoming. "Measure for Measure: A Prospective Study Comparing Quality Evaluation by Vignettes, Standardization and Chart Abstractions." *Journal of the American Medical Association*.
- Peabody, John, Omar Rahman, Paul Gertler, J. Mann, D. Farley, and G. Carter 1999. *Policy and Health: Implications for Development in Asia*. Cambridge, Mass: Cambridge University Press.
- RAND Corporation. 1998. "MatLab Health and Socioeconomic Survey Documentation." Santa Monica, Cal.
- Schultz, T. P., and A. Tansel. 1997. "Wage and Labor Supply Effects of Illness in Côte d'Ivoire and Ghana: Instrumental Variable Estimates for Days Disabled." *Journal of Development Economics* 53 (2): 251–86.
- Sen, Amartya. 1985. *Commodities and Capabilities*. Amsterdam: North-Holland.
- Sindelar, J., and D. Thomas. 1991. "Measurement of Child Health: Maternal Response Bias." Economic Growth Center Discussion Paper 663. Yale University, Department of Economics, New Haven, Conn.
- Stern, S. 1989. "Measuring the Effect of Disability on Labor Force Participation." *Journal of Human Resources* 24 (3): 361–95.
- Stewart, Al, and J.E. Ware, eds. 1992. *Measuring Functioning and Well-Being: The Medical Outcomes Approach*. Durham, N.C.: Duke University Press.
- Strauss, John. 1986. "Does Better Nutrition Raise Farm Productivity?" *Journal of Political Economy* 94 (2): 297–320.
- Strauss, John, and Duncan Thomas. 1995. "Human Resources: Empirical Modeling of Household and Family Decisions." In J. Behrman and T. N. Srinivasan, eds., *Handbook of Development Economics*. Amsterdam: North Holland.
- Strauss, J., P. Gertler, O. Rahman, and K. Fox. 1993. "Gender and Life-Cycle Differentials in the Patterns and Determinants of Adult Health." *Journal of Human Resources* 28 (4): 791–837.
- Wallace, R., and A. Herzog. 1995. "An Overview of Health Status Measures in the Health and Retirement Survey." *Journal of Human Resources* 30 (supplement): S84–S107.
- Ware, J., A. Davies-Avery, and R. Brook. 1980. *Conceptualization and Measurement of Health Status for Adults in the Health Insurance Study: Vol. IV, Analysis of Relationships Among Health Status Measures*. Report R-1987/6-HEW. RAND Corporation, Santa Monica, Cal.
- World Bank. 1993. *World Development Report: Investing in Health*. New York: Oxford University Press.

# 9 Employment

*Julie Anderson Schaffner*

Most households in developing countries earn most of their incomes from the productive employment of their members. Thus the employment module in a multitopic household survey collects information that is crucial for diagnosing the causes of poverty and inequality—and for guiding policymakers in their attempts to improve living standards.

The first section of this chapter provides an overview of the key policy issues that can be addressed with data from the employment module. The second section discusses the specific data needed to address these policy issues. The third section introduces three prototype designs of the employment module—a short, standard, and expanded version. (All of these versions are presented in Volume 3.) The fourth section is a list of annotations to the three draft modules.

## **Policy Concerns**

The employment module collects information that describes four main sets of labor market outcomes: employment and unemployment, earnings from wage employment, on-the-job training, and conditions of employment. This section highlights the policy issues that can be illuminated by the description and analysis of these outcomes.

### *Employment and Unemployment*

Policymakers in developing countries need accurate, up-to-date information on who works, the type of work they do, and the extent of unemployment, in order to answer important policy questions such as:

- What share of the working-age population has gainful employment?
- Which productive sectors of the economy employ large fractions of the work force?
- What is the relationship between employment and poverty, and what types of work are most common among the poor?
- How do work activities differ across region, age, sex, and ethnic or racial group?
- How common is child labor and what is its relationship to poverty?
- Which workers are employed in sectors of the economy that are directly affected by government rules and regulations such as minimum wage laws?
- Who participates in rural public works programs?
- Who works in the economic sectors that are most likely to be affected by changes in the international economy?
- Who is looking for work, how long have they been unemployed, and what are they doing to find employment?

Without a basic understanding of the nature of employment and unemployment, policymakers in a country will not be able to judge the importance of specific policy concerns. This subsection shows how

the employment module of multitopic surveys like the LSMS surveys can be used to answer a variety of important questions about the nature of employment and unemployment.

**LABOR FORCE PARTICIPATION.** Determining whether each survey respondent of working age is employed, unemployed, or “out of the labor force” (that is, neither employed nor seeking employment) makes it possible to estimate the size of the labor force and the size of the unemployment problem. Finding out why individuals are out of the labor force—for example, because they are caring for children or attending school or because they are handicapped—can also be useful for identifying and measuring which groups are most likely to be drawn into the labor force as economic conditions change. When questions about labor force participation are included in the employment module of a household survey, analysts can link the answers to these questions with data from the household roster and consumption modules to answer such diagnostic questions as: Are poor households poor because they have few able-bodied workers or because their workers earn low incomes? Are the nonworkers in poor households actively looking for work? If those who are currently unemployed were to receive unemployment assistance, would this have a greater effect on poor or nonpoor households?

**SECTOR CHOICES.** At least as useful as measuring whether individuals are employed or unemployed is identifying which individuals and households are involved in particular economic activities. Asking questions about the work people do and the employers for whom they work is relatively easy (compared to measuring wages and income), and such data make it possible to answer a wide range of diagnostic questions pertaining to the size of specific employment sectors. Even when the survey is administered only once, data on sector choices make it possible to answer such questions as: How large are the groups likely to be hit hardest by a decline in agricultural prices? How large is the group not covered by labor standards legislation and how large is the group outside the scope of payroll taxation? Do individuals in public sector employment have higher living standards than individuals with similar education in the private sector (which would raise the possibility that public sector wages could be cut without loss of employees)? How successful are rural

public works programs in targeting the poor? How many people are likely to be directly affected by an increase in the minimum wage, and would the beneficiaries be from the poorest households?

When the surveys are undertaken in more than one year, they illuminate how economic activities are changing over time, making it possible to answer additional questions such as: Do structural adjustment policies appear to be successful in shifting employment from import-competing to export sectors without greatly increasing unemployment? Did many households shift their economic activities into the informal sector after the income tax rate was increased? Is unionization rising or declining? Answering such questions can be an important first step in analyzing the costs and benefits of policy changes.<sup>1</sup>

When data are available on a particular household's economic activities at more than one point in time (either because a survey was fielded more than once to produce panel data or because it asked respondents to answer retrospective questions), it is possible to answer more “dynamic” questions about labor market flexibility. For example, when adjustment policies cause the distribution of workers across sectors to change, do workers from declining sectors move into growing sectors or do they move out of employment while new entrants to the labor market take up the new jobs in the growing sectors? When workers lose their public sector jobs, what sorts of alternative employment do they find?

**HOURS OF WORK.** Finding out not only whether individuals work but also how many hours they work helps quantify the intensity of participation in various sectors, and sheds additional light on the nature of poverty. It allows answers to such questions as: How high is the rate of labor utilization in the economy? For how many households engaged in agriculture have nonagricultural activities come to occupy a large share of their time? Are poor households poor because they find few hours of work, or are the poor already working long hours?

**JOB SEARCH METHODS AND THE DURATION OF UNEMPLOYMENT.** Asking the unemployed about the ways in which they search for jobs (for example, by reading newspapers, visiting potential employers, visiting public job service offices, or talking to friends) makes it possible to quantify the use of various search methods and helps policymakers determine the

potential for getting information to the unemployed through newspapers or job services. Such data allow analysts to answer the question: Which sorts of job referral programs, if any, reach the largest numbers of people with information about new and better jobs?

Asking the unemployed how long they have been unemployed allows researchers to assess the relative importance of long-term and short-term unemployment. With a large enough set of data on the duration of unemployment, a researcher can answer questions such as: How long is the typical period of unemployment? What characteristics do the long-term unemployed tend to have?

**UNDEREMPLOYMENT AND ON-THE-JOB SEARCH.** Some workers may have jobs that give them too few hours of work, make poor use of their skills, or earn them incomes that are lower than some subsistence level. Household survey data can be used to measure at least two sorts of “underemployment”: employment in jobs that do not yield earnings above some predetermined level and the search by employed workers for additional or alternative jobs. Data on underemployment as defined by income are useful for answering the same questions that are raised about poverty. Data on underemployment as defined by the search for additional or alternative work are much harder to interpret, but nonetheless are used in some countries as broad indicators of labor market conditions. It can also be useful to identify the job search methods used by workers who are looking for additional or alternative jobs. Such data might help answer questions such as: What is the best way for the government to disseminate information that will prompt workers to move out of declining sectors and into growing sectors?

#### *Earnings from Wage Employment*

Income levels are a key determinant of a household’s living standards and poverty status. In developing countries, income from work activities is by far the most important source of income for most households. An important component of work activities in most developing countries is wage employment, in which people work outside their homes for firms, the government, or other individuals, in exchange for payments in cash or kind. Thus policymakers need to know how much income workers earn from such activities. It is especially useful to know the wage (or, more precisely, the “average hourly earnings”) that

workers face. Wages provide a good description both of how attractive jobs are to workers and of how costly labor is to employers. Information on earnings from wage employment and hours worked in wage employment can be gathered in the employment module of multitopic household surveys like the LSMS surveys. When used to produce measures of wages, this information can be used to address many policy issues, including:

- How high are average wages in various sectors of the economy?
- How skewed is the distribution of wages?
- How are wage levels and the distribution of wages changing as the economy opens to international trade, or as the economy undergoes structural adjustment?

When combined with information from other modules on nonlabor income and income earned in agricultural and nonagricultural household enterprises, this information can also be used to address the questions raised in Chapter 17 related to the level and distribution of total income.

It is not possible to begin to answer such questions without data on earnings from wage employment. Such data are not, however, easy to collect. The second section of this chapter will discuss in detail the steps that can be taken to increase the accuracy with which data on income from wage employment can be collected.

Having data from a single survey makes it possible to answer questions about the wage employment sector such as: How attractive to workers are the average jobs in the wage employment sector of the economy? How unequally distributed are wages in this sector? In addition, combining these data with data from other modules makes it possible to analyze which workers receive high and low wages. (This is discussed further below, in the subsection on “Analyzing the Determinants of Labor Market Outcomes.”) When wage data are available from a series of cross-sectional surveys, this can show how the level and distribution of wages is evolving over time. An especially interesting question in recent years has been: How have the level and distribution of wages changed in the wake of reforms to open the economy to international trade and capital flows? When a panel of wage data are available, it becomes possible to address more detailed questions such as: How frequently do high-wage earners experience dramatic reductions in their wages or

do low-wage earners experience dramatic increases in their wages? Do workers who lose jobs in sectors forced into decline by economic reforms find new jobs at higher wages or at lower wages?

Though many previous surveys (both LSMS and others) have attempted to collect information on earnings from self-employment activities in the employment module, this is unlikely to produce adequate income data, for reasons discussed in Chapters 18 and 19 on agriculture and nonagricultural household enterprises.

### *On-the-Job Training*

Many workers improve their skills and acquire new skills by participating in on-the-job training. Skills thus acquired may be important for making workers more productive and causing their wages and living standards to rise. Yet so far little information is available on such training in developing countries. A variety of data about on-the-job training may be collected in the employment module of multitopic household surveys, allowing policymakers to address such questions as:

- How prevalent is on-the-job training, and in which sectors do workers find the most opportunities for such training?
- How does the prevalence of on-the-job training change when the economy opens to international trade or when the government attempts to subsidize training?
- Do low rates of on-the-job training suggest that the economy faces obstacles to entry into modern manufacturing and service activities (in which on-the-job training is important)?

Even when collected in a single survey, employment module data on training can shed light on whether workers have received or are currently receiving some training, what kind of training they have received (for example, learning by doing, learning informally from others, on-site formal training, off-site formal training, or apprenticeship), how long the training lasted, in which occupations they have received training, how much they paid for their apprenticeship, and what sorts of employers provide training. With this information it becomes possible to answer such questions as: How large is the stock of human capital that was acquired through on-the-job training? What sorts of training appear to be most important? In which sectors and in which sorts of jobs is on-the-job training most prevalent? In which sec-

tors do on-the-job training rates in the country studied differ most from on-the-job training rates in developed countries? When questions about training are asked in a series of cross-sectional surveys, the evolution of training can be traced over time. This may be especially useful in countries that are becoming increasingly involved in the global economy or countries in which the government has attempted to encourage on-the-job training through subsidy or tax rebate programs. When panel data on training are available, it is possible to assess differences in the career paths of workers who do and do not acquire training.

Though a household-based multitopic survey like the LSMS can shed light on these questions, it cannot provide all the information needed for a thorough evaluation of training. For example, only employers could provide detailed answers to questions about the costs of on-the-job training.

### *Conditions of Employment*

Wages are not the only characteristic of work that is of interest to policymakers. Policymakers are also concerned about other features of employment such as sick leave, pensions, health insurance, and job security. Governments would like to know the extent to which such nonwage benefits exist and what policies the governments can adopt to encourage these benefits. Governments often attempt to legislate these benefits, but in developing countries such laws are difficult to enforce and, in some cases, may have unintended negative consequences on workers. Information collected in the employment module of a multitopic household survey allows policymakers to address such questions as:

- How well enforced are regulations regarding nonwage benefits?
- When regulations regarding a particular nonwage benefit change, how does the prevalence of that and other nonwage benefits change in response?
- How stable are jobs in various sectors of the economy, and how is job stability affected by job security legislation?

**NONWAGE BENEFITS AND WORKING CONDITIONS.** Finding out about workers' nonwage benefits (such as paid holidays, paid sick leave, and pensions), working conditions, the costs of their commute to work, and the location of their homes (whether at the work site or elsewhere) can be useful for answering such ques-

tions as: To what extent do employers comply with laws requiring them to provide specific nonwage benefits? If nonwage job attributes are taken into account, is there more or less inequality in worker well-being than when only wages are taken into account?

When such data are available for two or more time periods, they are also useful for answering questions such as: How are living standards, defined more broadly than mere measures of wages or incomes, changing over time? When regulations increasing specific nonwage benefits take effect, does the incidence of these benefits increase, and do the wages or other benefits of the recipients appear to decline in compensation for the increase in nonwage benefits?

**JOB TENURE AND TURNOVER.** One important feature of employment, job stability, is difficult to observe directly, but analysts can draw inferences about it by asking workers how long they have worked for their current employer. Analyzing data on job tenure can produce answers to such questions as: How long do jobs in various sectors of the economy typically last? How does job tenure in the country studied compare with job tenure in other developed countries? (See Hall 1982 for a method involving a single cross-section dataset; see Schaffner 1999 for methods involving single and repeated cross-sections.) When data are available from a series of cross-sectional surveys, it becomes possible to answer the question: What has been the apparent effect of changes in job security legislation on the typical length of a job?

In order to find out what the labor market is like for workers, it is useful to know not only how quickly jobs end but also how they end. Clearly, the labor market is very different if most workers that lose their jobs are unemployed before finding other jobs than if most workers quit their jobs in order to take up better jobs that have become available.

**TYPES OF EMPLOYMENT CONTRACTS.** In many countries the law specifies what kinds of employment contract an employer can offer; each allowed contract must satisfy certain requirements regarding nonwage benefits and job security. For example, workers hired under contracts of “indefinite duration” may be entitled to a variety of nonwage benefits and to large severance payments if they are dismissed without legally defined “just cause.” Workers hired under contracts of “fixed duration” may be entitled to no such benefits or

severance payments, but the legal use of such contracts is often severely restricted to very short durations, to limited shares of firms’ work forces, and to specific sectors. In an attempt to assess the role that such laws play in labor markets, it is useful to ask: How many workers have explicit or written employment contracts, and how many workers remain outside the scope of the legislation? Which legal contract types predominate in which sectors? Do workers with written contracts have better mandated nonwage benefits than workers without written contracts and to what degree? Do patterns of job turnover differ between workers with explicit, written contracts and workers without such contracts?

Data on types of legal contracts can be collected in the employment module of LSMS-type surveys, allowing researchers to answer these questions. When such data are collected in two or more years, it becomes possible to answer such questions as: Did a reduction in the severance payments required upon dismissal from contracts of indefinite duration appear to increase the use of such contracts?

#### *Analyzing the Determinants of Labor Market Outcomes*

Descriptive information on the existing state of the labor market, such as that discussed above, is useful for policy analysis. Policymakers would, however, like to know even more about how policies are likely to affect employment, sector choice, wages, training, and other labor market outcomes. This usually requires estimating econometric relationships between labor market outcomes, on the one hand, and a variety of individual, household, employer, job, and community characteristics, on the other. If a reasonably complete list of potential determinants is included in the estimated relationship, the estimates allow the policymaker to assess what would happen to wages, sector choice, and other outcomes if the government intervened to change one determinant while all others remained the same (see discussion of multiple regression in Chapter 26 on econometrics). Such analyses help policymakers answer questions such as:

- How big is the impact of expanded primary education, or of improved health care, on the incomes of the poor?
- Could the wages of some public sector workers be reduced (improving the government deficit) without losing these workers to private sector jobs?

- Does minimum wage legislation appear to help an elite among workers, while driving down wages for the rest?
- Should the government consider developing training loan programs or industry training consortia to promote on-the-job training?
- By how much is the income tax base likely to shrink (as workers leave the sectors in which income tax law is enforced) when the income tax rate rises?

This subsection will discuss the usefulness for policy analysis of including measures of a variety of determinants in relationships that study the determination of labor market outcomes. Where an empirical literature has emerged, one or two key references will be cited; in some areas the potential for useful empirical research remains untapped.

**INDIVIDUAL CHARACTERISTICS.** Estimating the effects on wages and other labor market outcomes of individual characteristics such as schooling (of different levels, types, and qualities), training, health, age, gender, and ethnicity helps policymakers answer the questions: How large are the returns (in the form of increased earnings) to investments in schooling, training, health, and nutrition? At which schooling level do investments have the highest returns? Which dimensions of school quality are most important for improving students' earnings potential? Do differences between the productive characteristics (such as number of years of schooling) of men and women explain male-female differences in average wages, labor force participation, training, and participation in high-wage sectors—or is there reason to worry about labor market discrimination against women? (Strauss and Thomas 1995 review much of the relevant literature.)

**CHARACTERISTICS OF PARENTS' HOUSEHOLD.** Estimating the effects on wages and other labor market outcomes of the education and occupations of the worker's parents makes it possible to answer such questions as: Does the worker have a better job and a higher level of education than his or her parents (Heckman and Hotz 1986)? Are better-educated parents able to provide their children with better "connections," thus increasing their children's prospects of finding high-paying jobs (Lam and Schoeni 1993)?

Collecting information that describes the financial constraints that the family may have been facing at

the time when the individual was entering the labor market allows study of the role of financial constraints in determining workers' career paths. Financial constraints may play an important role in determining career paths and lifetime earnings, because individuals sometimes need to sacrifice earnings in the early years of their careers to get on a career path that will bring them higher earnings later in life. Families with little disposable income or access to credit may not be able to make such an investment in a good career.<sup>2</sup> Having these data on families' financial constraints helps policymakers answer the question: What evidence is there that financial constraints early in individuals' working lives cause them to go into occupations with little wage growth and low lifetime earnings?

**EMPLOYER, WORKPLACE, AND JOB CHARACTERISTICS.**

The main reason policymakers are interested in knowing how the characteristics of a worker's employer, workplace, and job affect his or her wages is that labor markets in developing countries are often thought to be severely "segmented." Such segmentation would mean that even among workers with identical interests, abilities, and training, some get much higher wages and better benefits than others. Segmentation implies that labor is allocated inefficiently in the market (because too few workers are employed in the high-wage, high marginal product sectors) and that some earnings inequality is of an especially troubling sort, in which luck and connections play a large part in determining income. Determining whether labor markets are segmented is important because, as discussed in Bulow and Summers (1986), the best policies for welfare when labor markets function well may be very different from the policies that are best when labor markets are segmented.

Segmentation can take many forms. Segmentation can arise between public and private sector workers if the public sector sets wages higher than necessary to attract workers. Segmentation among geographic regions may arise if workers have difficulty acquiring accurate information about wages in other regions, or if moving between regions is costly. Segmentation between union and nonunion workers may arise if labor unions are able to obtain higher wages for their members through bargaining. If minimum wage legislation is imperfectly enforced, this may cause segmentation between workers at firms that comply and workers at firms that do not. Finally, some (larger,

more modern) employers may choose to pay higher wages than others, for example, to give their workers an incentive to work hard (Bulow and Summers 1986) or to reduce the frequency with which their workers quit and must be replaced (Stiglitz 1974).

Data in LSMS surveys can help identify the nature, severity, and causes of segmentation. Estimating the effects on wages of sector (public versus private), region, coverage by collective bargaining agreements, and coverage by minimum wage legislation—as well as employer size and sector—allow one to begin answering questions such as: Which public sector workers receive wages that are particularly out of line with private sector wages? Would improving transportation and communication networks reduce segmentation among regional labor markets? To what extent do labor unions drive their members' wages up relative to the wages of those nonunion workers who have comparable skills? How large are wage differentials caused by incomplete enforcement of minimum wage legislation? Do labor markets appear segmented between large, modern firms and other employers, for the more complex reasons discussed above? (See Schaffner 1998.)

Some questions about labor market segmentation are beyond the scope of multitopic household surveys like the LSMS. For example, probing more deeply into the causes of some sorts of labor market segmentation would require detailed information from employers about technology, supervision difficulties, training, and searches for job candidates.

**PREVIOUS LABOR FORCE STATUS AND LOCATION OF RESIDENCE.** If the probability of a worker obtaining a high-wage urban job varies depending on his or her status in the labor force or on where he or she lives, this can have important policy implications. For example, if the probability of getting a high-wage job is higher if one searches while unemployed than if one searches while working in a low-wage job, some workers may choose to be unemployed (and let their labor go to waste) even when low-wage jobs are available. This makes labor market segmentation (which underlies the existence of both high- and low-wage jobs for identical workers) more costly to the economy than it otherwise would be, and suggests the difficulty of reducing unemployment without first tackling the causes of labor market segmentation. Similarly, if the probability of obtaining a high-wage job is high-

er in urban areas (even in low-wage employment) than in rural areas, as in Harris and Todaro's famous 1970 model, rural-urban migration will tend to be excessive and policies to promote agricultural employment may be called for.

**TRAINING ON CURRENT JOB.** Training is of interest to policymakers not only as an outcome variable but also as a determinant of wages and sector choice. Estimating the effects of training on other labor market outcomes can show whether workers must accept lower wages to obtain jobs that give them training. If this is the case, financial constraints on workers from poor households might prevent these workers from obtaining training and better careers, and experimentation with training loans might be justified.

**TENURE IN JOB, OCCUPATION, AND INDUSTRY.** Do wages rise as workers' length of tenure on the current job rises? If so, there is reason to think that employers increase labor productivity by encouraging workers to remain in the same job for a long time. Increasing wages to keep workers may economize on training costs or provide the workers with an incentive to work hard. Unstable macroeconomic policies or other policies that make it difficult for workers and employers to reach long-term employment agreements may hinder such improvements in labor productivity.

Do wages rise with tenure in some industries, regardless of whether workers are still with the employer from whom they received training? If so, fear of poaching may make employers in that industry reluctant to provide training. However, encouraging the development of training consortia within those industries might overcome this disincentive to provide training. Do wages start lower and rise more rapidly with tenure in some occupations than in others? If so, borrowing constraints may prevent some workers from entering into occupations with better career potential for which they are qualified and motivated.

**COMMUNITY CHARACTERISTICS.** Many policy effects on wages and labor supply can be estimated by comparing outcomes across communities that have been affected in differing degrees by the policies of interest (see Chapter 13 on the community questionnaire).

Gathering data on community-based measures of infrastructure, prices that affect microenterprise profits, ease of transportation to major wage employing

regions, quality and cost of local schools, and availability and attractiveness of rural public works programs allows researchers to address such questions as: How big an impact do microenterprise support programs have on the choice to become self employed? To what extent would the construction of better roads affect commuting behavior? How would the incidence of child labor be affected by improvements in the quality of schools? How do changes in rural public works program eligibility and in program wages affect both program participation and allocation of participating households' labor to other wage employment and to household enterprise activities? (If these other activities drop off when participation in the program increases, the program's effect on household income is diluted.) As discussed in Chapter 23, repeated cross-section or panel data on the same communities might be preferred to single cross-section data for estimating these effects when programs are endogenously placed or migration is selective.

**INCOME TAX AND TRANSFER POLICIES.** An increase in the income tax causes the effective wage (the amount by which the worker's income increases when he or she works an additional hour) to fall, at least in sectors in which the income tax law is enforced. By reducing the effective wage in such sectors, an increase in the income tax may reduce work hours, cause workers to move to jobs in sectors in which the income tax law is not enforced, or reduce overall participation in the labor force. Such a tax increase has the same effect on labor supply as a decrease in the actual wages that causes the effective wage to change by the same amount. Similarly, many income transfer programs that aim to alleviate poverty have the same effects on labor supply and sector choice as do comparable changes in effective wages and household nonlabor income (see Chapter 11 for definitions of nonlabor income).<sup>3</sup>

Good estimates of the effects of wages and nonlabor income on labor supply and sector choice would allow policymakers to answer such questions as: When the income tax is increased, by how much does the tax base shrink and what happens to total tax revenue? When the income tax is increased, how much labor becomes less productive as workers leave the labor force or move to less productive sectors of the economy where the income tax law is not enforced? If a means-tested income transfer program were introduced as part of a new "social safety net," how much

labor would be rendered less productive as recipients reduced their labor supply or shifted into less productive sectors in which their earnings were "off the books" (and thus could not be used to reduce the size of their transfers)? Unfortunately, the estimation of labor supply responses to wage changes is complicated and often yields results that do not inspire confidence (Mroz 1987). For reasons discussed in the second section of this chapter, wage effects on labor supply and sector choice are best estimated in repeated cross-section or panel data. Even with such data, however, wage effects can be estimated only if one is willing to make very restrictive assumptions about how people make labor supply decisions. The necessity for such assumptions renders current methods especially unattractive for the study of developing economies.

### *Summary*

Box 9.1 summarizes the most important policy issues pertaining to labor markets in developing countries. The table indicates which issues can be addressed with just one set of data, which issues can be addressed with a series of data sets from several cross-sectional surveys, which issues need panel or retrospective data, and which issues are difficult to study effectively with data from LSMS-type surveys.

### **Data Needs**

This section discusses the implications of the research objectives listed in the first three sections of Box 9.1 for the design of multitopic household surveys.

### *Survey Design Issues*

Before discussing how to collect data to answer the specific policy issues that were raised in the first section, some general survey design choices must be discussed. These involve choice of reference periods, handling of multiple activities, data collection over time, and linking of the questions in the employment module to the questions in the household enterprise module.

**REFERENCE PERIODS.** Given the importance of agriculture in most developing countries, it is common for the economic activities of individuals and households to vary during a given year. As a result, the answers that respondents give to questions about whether or not they worked, the sectors in which they worked, and their average hourly earnings will depend upon which

**Box 9.1 Labor Market Policy Issues and Data from Multitopic Household Surveys**

*Issues that can be addressed with a single round of survey data*

1. Measuring labor force participation, employment, and unemployment, and reasons for being out of the labor force.
2. Measuring participation in various sectors of the economy and types of economic activity.
3. Identifying the labor market status of the poor and the sectors in which they are employed.
4. Measuring hours of work, both in total and by sector.
5. Measuring the duration of periods of unemployment and what search methods are used by the unemployed.
6. Measuring underemployment and on-the-job search methods.
7. Measuring the mean and distribution of average hourly earnings among the employed.
8. Estimating the mean and distribution of average hourly earnings among the self-employed (using data from the household enterprise and the agriculture module rather than the employment module).
9. Measuring the mean and distribution of total annual earnings from wage employment.
10. Measuring the mean and distribution of total annual earnings from self-employment (again using data from the household enterprise and agriculture modules rather than the employment module).
11. Measuring current and past participation in formal training.
12. Measuring current and past participation in informal training.
13. Identifying which employers provide and which workers receive formal and informal training; describing workers' occupational experience.
14. Measuring the incidence of various nonwage benefits and describing working conditions.
15. Measuring the distribution of jobs by the length of the job and studying the nature of job turnover.
16. Measuring the incidence of various types of employment contracts.
17. Analyzing the effects of schooling, health, and nutrition on wages and other labor market outcomes.
18. Analyzing the effects of training and work experience on wages and other labor market outcomes.
19. Analyzing the relationship of contractual relationships to wages and other labor market outcomes.
20. Analyzing the likely importance of labor market discrimination by estimating wage differentials by race or gender after controlling for such complicating factors as skills and working conditions.
21. Analyzing the degree of intergenerational labor market and earnings mobility.

22. Analyzing the effects of early-life financial constraints on career paths.
23. Assessing the likely importance of labor market segmentation across regions, sectors, and employer types by estimating wage differentials by region, sector, and employer type after controlling for such factors as skills and working conditions.
24. Assessing the likely importance of differences among communities in infrastructure prices and location on labor market outcomes.

*Issues that can be addressed using repeated cross-sectional, panel, or retrospective survey data*

25. Describing changes over time, possibly associated with policy changes, in labor force participation in various sectors of the economy and participation in training (repeated cross-sectional, panel, or retrospective data).
26. Describing changes over time, possibly associated with policy changes, in the level and distribution of average hourly earnings in wage employment (repeated cross-sectional or panel data).
27. Describing changes over time, possibly associated with policy changes, in the incidence of nonwage benefits (repeated cross-sectional, panel, or retrospective data).
28. Describing changes over time, possibly associated with policy changes, in the incidence of various employment contract types (repeated cross-sectional, panel, or retrospective data).
29. Measuring how expected wages affect labor force participation and sectoral choice in simple models (repeated cross-sectional or panel data).
30. Assessing the degree of occupational and sectoral mobility in the labor market (panel or retrospective data).
31. Measuring the relative success of various job search methods in obtaining high-wage jobs (panel or retrospective data).
32. Assessing income mobility (panel data).
33. Identifying winners and losers associated with policy changes (panel data).
34. Assessing how career paths differ between workers who acquire training and workers who do not. Also, how career paths differ between workers who obtain "good" jobs early in life and workers who do not (panel or retrospective data).
35. Assessing the effects of risk and shocks on labor market choices (panel data).
36. Assessing the relative success of workers' searches for high-wage jobs by their labor force status and by the location of their residence in order to shed light on the

*(Box continues on next page.)*

**Box 9.1 Labor Market Policy Issues and Data from Multitopic Household Surveys (continued)**

causes of open unemployment and on whether rural-urban migration is excessive (retrospective data or a panel that follows migrants).

*Issues that are difficult to address with multitopic household survey data*

- 37. Measuring the costs of on-the-job training (for which an employer survey is needed, possibly in conjunction with a household survey, so that workers' wages can be examined in light of both worker and employer characteristics).
- 38. Quantifying nonwage benefits (which would require a more focused survey).

*Source: Author's summary.*

- 39. Examining in detail how wage levels are related to employers' technological conditions (which again would require an employer survey, possibly in conjunction with a household survey).
- 40. Measuring the effects of temporary and permanent wage changes on labor force participation, hours, and sector choice in dynamic models that acknowledge the potential effect of today's choices on tomorrow's opportunities (which would require panel data of sufficient length, and methodological development, as discussed in the second section of this chapter).

reference period is given in the question. In other words, their answers will depend upon whether the questions refer to the previous week, the previous month, or the previous year. Thus choosing the most appropriate reference period is of great importance. The survey designers must first decide which reference periods will yield data of the most use to analysts. Having made this decision, they may elect to use different reference periods in some particular questions—which is acceptable provided that the module also contains questions that elicit data required for converting measures into the reference periods of most use to analysts.

For many analytical purposes, the reference period of greatest interest is one full year (either the previous 12 months or the previous calendar year). For example, when attempting to identify households who depend on agriculture for a large fraction of their living, it is important to avoid asking respondents questions about their sector of employment that refer to the previous day or to the previous 7 days; if the interview occurs during the agricultural slack season, respondents might report no involvement in agriculture, even though they may devote a large portion of their time to agriculture each year. Similarly, if households know that their earnings fluctuate over the year and if they are able to save enough during their high-income seasons to cover their consumption during their low-income seasons, annual earnings data would be a much better indicator of their living standards than data on their earnings during the previous 7 days.

Nevertheless, analysts may sometimes prefer to use the 7-day reference period for at least two reasons.

First, most labor force surveys around the world use a 7-day reference period to collect data on labor force participation, employment, and unemployment. Thus, if analysts want to compare labor market performance in the country of the survey to performance in other countries, it is essential for the survey to use a 7-day reference period. Second, questions that use shorter reference periods tend to elicit more accurate responses than questions that use longer reference periods, both because they do not stretch respondents' memories and because it is easier to design brief sequences of questions when a shorter recall period is used. The 7-day reference period is usually the most appropriate short reference period to use, because it captures the effects of weekly days off on hours and earnings.

This chapter recommends that survey designers aim to use both the 7-day and the 12-month reference period, as has been common in past LSMS-type surveys. However, unlike in many earlier LSMS-type surveys, it is recommended that the questions that use the 12-month reference period should be less ambitious than those that use the 7-day reference period, because it is harder for respondents to recall details of their work experience from up to 12 months before than to recall equivalent details from the previous week.

The costs of using a 7-day reference period can be minimized if random samples of households are interviewed in each month or season of the year, as has been done in many past LSMS-type surveys. It is strongly recommended that survey organizers maintain this feature, and the rest of this discussion assumes that this will be the case. When this intrayear sample design is used, the answers to the questions using the

7-day reference period can be used to answer some questions about annual earnings. For example, the sample average of 7-day reference period earnings for a group with a particular level of education is an estimate of the average (across weeks of the year as well as across individuals) of earnings for this educational group in the population as a whole. Therefore, multiplying the result by 52 provides an estimate of average annual earnings in that educational group. Repeating this exercise for groups of people with other levels of educational attainment will show whether and how annual earnings differ across workers with different levels of education.

Although interviewing random samples of households in each month of a year reduces the costs of using the 7-day reference period, it does not eliminate these costs. Questions that use a 7-day reference period do not yield good estimates of the share of households that ever participated in a particular activity during a specific year, even when they are administered to households over the course of 12 months. Nor do these questions yield sufficient data to make possible a complete analysis of wage or income distribution or of poverty (measured using income rather than consumption data). Although the data that these questions yield make it possible to study how average annual earnings vary across groups with different observable characteristics (like schooling levels and locations), it is not possible to use 7-day reference period earnings data to study inequality in annual earnings within groups. The problem is that when annual earnings are studied by examining 7-day reference period earnings of similar (but not identical) households in different months of the year, it is impossible for analysts to tell if the individuals and households that have high incomes in one month also have high incomes in other months. Therefore, it is recommended that the module collect data using 12-month as well as 7-day reference periods.

Even though the objective of the draft employment modules introduced in the third section of this chapter (and presented in Volume 3) is to measure labor market outcomes using 7-day and 12-month reference periods, respondents are allowed to report their cash and in-kind earnings in whatever reference period they like (from one hour to one year). This is because earnings are especially difficult for respondents to report accurately, and it is hoped that by allowing respondents to respond in the way most nat-

ural to them, interview time will be reduced and data yielded will be more accurate. In order to turn these responses on earnings into average hourly, weekly, and annual earnings figures that are comparable across individuals, it is necessary to include additional questions in the module. To construct average hourly earnings figures, reported earnings in cash, food, and other forms of remuneration should be divided by the number of hours worked during the reference period that the respondent chose. Adding together these average hourly earnings in cash, food, and other forms of remuneration will produce a measure of the respondent's total average hourly earnings for a particular job. This figure must then be multiplied by the number of hours worked during the previous 7-day or 12-month period to produce measures of earnings that are comparable across individuals or households.

There are two ways to estimate the number of hours worked during the reference period chosen by the respondent. First, the interviewer can ask respondents directly how many hours they worked during the time period for which they reported their earnings. Second, the interviewer can ask respondents a set of questions about their average daily hours, average weekly hours, and weeks worked per year, which will allow analysts to construct rough estimates of the hours worked by respondents in almost any reference period. The first approach is likely to yield more accurate estimates of average hourly earnings but would be difficult to use to calculate every component of earnings on an hourly basis. Because respondents are allowed to choose a different reference period for each component of their earnings, it would be necessary to follow each question about one component of the respondent's earnings with a question about the number of hours he or she worked during the reference period that the respondent chose for that component. The second approach is less accurate, but asking one short set of questions about hours makes it possible to convert all of the respondent's earnings components into an hourly figure.

It is suggested that, for the purpose of converting reported cash wages into an hourly figure, the interviewer should ask respondents directly about the number of hours that they worked during their chosen reference period. To convert the in-kind components of respondents' wages into an hourly figure, however, the interviewer should ask them a short set of questions about the number of hours worked per

day, hours worked per week, and weeks worked per year, from which analysts can construct rough estimates of the number of hours worked in any reference period. This requires adding only one more question (on hours worked in the reference period chosen for reporting cash earnings) than has been included in most previous LSMS surveys, and it is likely to increase the accuracy of estimates of average hourly earnings substantially, because cash earnings are the most important component of earnings in most wage employment.

**MULTIPLE ACTIVITIES.** In most developing countries it is common for households to be involved on an ongoing basis with more than one economic activity. Not only might different household members work in different activities, and some or all household members work in different activities in different seasons of the year, but many individuals are involved in more than one economic activity in any given week. Thus, whether the 7-day or 12-month reference period is used, survey designers need to be aware that these individuals are involved in multiple activities. This has two implications for survey design.

First, because workers may perform a certain type of work for more than one employer (or for an employer and in a household enterprise), survey designers must be careful to distinguish between questions about occupations (in other words, work of a certain type performed for any employer or enterprise) and jobs (in other words, work of a certain type performed for a particular employer or enterprise). For example, if the questionnaire asks respondents about the number of hours that they spend working in their main occupations (possibly for more than one employer) but asks them about their earnings only from their main employer, this may result in average hourly earnings being greatly understated for individuals who perform their main work for several employers. The wording in previous LSMS surveys has not been very clear in distinguishing between jobs and occupations. The draft modules introduced in the third section of this chapter attempt to be more careful about making this distinction. Because it is important (for the analytical purposes discussed above) to associate average hourly earnings with particular types of employers, the module specifies earnings and hours from the “main employer” in a particular occupation, which should make it possible to measure respondents’ average

hourly earnings from a particular employer. However, because it may also be important to measure respondents’ total earnings from all employers, the draft module also asks respondents about the total number of hours they spent working in the occupation (for any employer). Collecting this information makes it possible to estimate respondents’ total earnings from all their employers in that occupation by multiplying their total number of hours by their average hourly earnings from the main employer.

Second, survey designers must allow respondents to give information about more than one occupation. They must decide the number of occupations in each reference period on which they will collect various amounts of information, and define a rule for choosing which activities should be described in detail when individuals are involved in more activities than are allowed in the detailed questions. Most previous LSMS surveys have allowed respondents to choose what they consider their “main work” and their “secondary work.” Then the respondents have been asked a complete set of questions about their main work, a shorter set of questions about their secondary work, and only a single summary earnings question about any other work. The surveys have typically asked respondents about the existence of any secondary work only after asking them a long list of questions about their main work, which may have made respondents reluctant to report another job on which they might have had to answer a battery of questions.

In contrast, the draft module begins by asking respondents to describe briefly all of the occupations in which they have been engaged during the previous 7 days and 12 months. Then the interviewer picks the two most “important” activities in each reference period based on the number of hours that the respondent worked in those occupations. More detailed questions are asked about the work that the respondent has done for the main employer in each of these two activities. It is hoped that this approach will reduce any inaccuracies that might be introduced by respondents’ reluctance to report secondary and tertiary activities and will increase comparability across respondents by making the definition of “main” and “secondary” jobs more concrete.<sup>4</sup>

**COLLECTING REPEATED CROSS-SECTIONAL, PANEL, AND RETROSPECTIVE DATA.** Collecting data on employment variations over time is of great value to analysts

and ultimately to policymakers. Collecting these data can be done in three ways. An original survey can be repeated using the same sample (panel data) or a new sample representative of the population at that later date (repeated cross-sectional data). Or a survey fielded only once can include questions about respondents' economic activities at some time in the past (for example, 5 years ago) as well as in the past 7 days or 12 months (retrospective).

To address issues 25–28 in Box 9.1, which pertain to broad changes over time in various labor market outcomes, it is necessary to have data from more than one time period using any of these methods. When a “before” survey has been conducted, the best way to examine these issues is to conduct repeated cross-sectional surveys. Retrospective data are likely to be less accurate than data gathered from an original cross-sectional sample, because respondents may have difficulty remembering detailed labor market outcomes over periods of one or more years. Panel data may provide a poorer picture (than would repeated cross-sectional data) of how labor market outcomes are changing over time in the population, because after the first sample, panel data are unlikely to be representative of the population as a result of sample attrition (although Chapter 23 discusses how this problem can be mitigated). When no baseline cross-sectional survey was fielded at the beginning of the period in question, retrospective data are the only option.

While the need for data from various time periods is clear in the case of issues 25–28, it requires some explanation in the case of issue 29, which pertains to the effect of wages on labor force participation, hours, and sector choice. Estimating such effects is unfortunately quite complicated. A key problem is that the analyst would like to estimate the effect on labor force participation—and participation in various sectors of the economy—of potential wages in each sector of the economy in which the worker might choose to participate. For example, the analyst would like to estimate the effect on participation in the private sector of the wages the worker could expect to receive in both private and public sectors. That would require measuring potential public sector wages even for individuals who chose to work in the private sector or not at all. Because it is nearly impossible to collect such information, researchers must attempt to estimate the wages workers could hope to obtain in various sectors, and then include these estimated or predicted wages as explanatory variables.

Estimating the effects of predicted wages on labor supply and sector choice is possible only when the predicted wage variable used on the right-hand side of the labor supply and sector choice equations is a function of some observed variables that affect labor supply and sector choice only through their effect on wages. That is, the dataset must contain some variables that affect wages but do not affect labor supply by changing workers' “preferences” toward staying home to perform household tasks or toward avoiding certain kinds of working conditions.<sup>5</sup> Unfortunately, most individual, household, and community characteristics that influence wages might also influence preferences. If workers' preferences do not change much from year to year, year indicators (in repeated cross-section or panel data) might fulfill the requirement of being important determinants of wages that do not affect labor supply and sector choice in any other way. Thus repeated cross-section or panel data might be useful for estimating wage effects, but only if a sufficient number of years of data, and thus a sufficient amount of independent wage variation, is available.<sup>6</sup>

To address issues 30–36 in Box 9.1 it is necessary to follow the same individuals over time either by building a panel of data on the same respondents or by collecting retrospective data. Collecting retrospective data has several advantages over collecting panel data for studying these questions. First, retrospective data are based on samples that are not subject to attrition—a factor that often renders panel datasets unrepresentative of the population. Second, retrospective datasets tend to contain fewer spurious changes over time (induced by measurement errors) than panel data.<sup>7</sup> Finally, it is possible to use retrospective data to study changes over periods for which no “before” data are available. As mentioned above, the major disadvantage of retrospective data is that they tend to be full of recall errors, although these are probably less severe in qualitative measures (such as answers on the sector of activity) than in quantitative measures (such as answers on earnings). Thus the draft module contains retrospective questions about qualitative outcomes (pertaining to the nature of activities) but not quantitative outcomes (such as hours and earnings). Panel data must be collected to study the changes in hours and earnings experienced by individuals over time.<sup>8</sup>

Survey designers should decide the best time interval to span with panel or retrospective data according to the main measurement objective of the

survey. For describing the short-term effects of a structural adjustment program, an interval of one year may be appropriate. For studying long-term development concerns over periods during which there have been no significant changes in government policy or in the country's economic circumstances, five-year intervals may be preferable. The draft module contains five-year retrospective questions, which can easily be modified to cover a shorter period.

**LINKS TO THE HOUSEHOLD ENTERPRISE MODULE.** In most past LSMS surveys, analysts could study the profits of nonagricultural household enterprises only after linking data from the household enterprise module with data from the employment module for the same households, because data on hours worked in household enterprise activities were collected only in the employment module. However, making this link was extremely difficult (see Chapter 18). The draft modules introduced by the third section of this chapter (and presented in Volume 3) include questions that facilitate making that link, but these difficulties can also be reduced by including questions about work in household enterprise activities in both the employment and household enterprises modules. Questions on the supply of labor should be asked in both modules, because the labor supply measures needed for employment analysis and for analysis of household enterprises are quite different. For comparability with labor force surveys from around the world, and for other reasons, labor supply measures for employment analysis must refer to 7-day (and 12-month) reference periods, and should be asked of the individual who performed the work. For reasons discussed in Chapter 18, labor supply questions used for analysis of household enterprises should refer to 2-week (and 12-month) reference periods, and should be asked of the manager of the enterprise.

Therefore, the employment module must be preceded by a brief section that identifies all of the household enterprises in the household. To ensure that the survey collects earnings data for all jobs, the questions used in that introductory section must be phrased so as to prompt the respondent to include any activities for which earnings data are not collected in the employment module. For example, they should be phrased so as to include the activities of itinerant carpenters, whose earnings are better measured in the household enterprise module but who may not think of themselves as operating an "enterprise."

To make explicit links between employment module data on labor supply and household enterprise activities, questions can be included in the employment module that ask respondents who have characterized themselves as self-employed to identify in which household enterprise (reported in the brief household enterprise roster module) this work was performed. This is likely to reveal some additional self-employment activities that should be included in the household enterprise roster, and thus serve as a check on the accuracy of the household enterprise roster, as well as a means of linking the modules. (This places an extra burden on the interviewer, who must do some cross-checking during the interview; the cross-checking will become easier as questionnaires become computerized.)

#### *Designing Questions about Specific Labor Market Outcomes*

This subsection addresses how to design questions that elicit complete and accurate information about the four main sets of labor market outcomes that were discussed in the first section: employment and unemployment, earnings from employment, on-the-job training, and conditions of employment.

**EMPLOYMENT AND UNEMPLOYMENT.** According to International Labor Organization guidelines, individuals are employed if they worked for at least one hour during the previous 7 days, either for a wage, in a household enterprise, or as an unpaid apprentice or trainee. Individuals also count as employed if they did not work during the previous 7 days but had a permanent job from which they were temporarily absent. Individuals are unemployed if they were not employed and were looking for work during the previous 7 days. They are participating in the labor force if they were either employed or unemployed during the previous 7 days.

Measuring employment is especially difficult in developing countries, where many income-generating activities are performed outside of markets and may not be thought of by respondents as "work" or "employment." (For a detailed discussion see Hussmanns and others 1990.) In an effort to elicit accurate data on labor force participation, the proposed employment modules include a sequence of questions that ask about three kinds of work (own farm, nonfarm household enterprise, and wage

employment) as well as about job searches and temporary absences from permanent jobs. The questions also provide examples with the aim of helping respondents understand the range of activities to which the interviewer is referring. In an effort to assess the extent of underemployment and on-the-job searches, the prototype module contains (as optional extensions) questions on whether respondents with jobs were looking for additional or replacement work during the previous 7 days. The wide range of sector distinctions of analytical interest are discussed above. For detailed notes on specific questions see the fourth section of this chapter.

Care must also be taken in measuring hours of work. With the aim of allowing researchers to construct measures of hours for any reference period a respondent might use for earnings reports, previous living standards surveys have asked respondents about days worked in the last week, hours worked per day in the past week, weeks worked in the past year, and usual hours worked per week during weeks worked in the past year. U.S. validation studies (which compare workers' reports about hours and earnings in household surveys with employer records) suggest that one should try to avoid the use of questions about "usual" or "average" hours, because (even in the less seasonal U.S. environment) responses to such questions contain a great deal of error as measures of the average hours researchers hope to measure (Rodgers, Brown, and Duncan 1993). In light of this, the draft employment modules replace the question on average hours worked per day in the past 7 days with a question on total hours worked in the past 7 days. Researchers can calculate average daily hours themselves if they wish. Unfortunately, it is difficult to avoid the use of "usual" weekly hours questions when attempting to measure 12-month labor supply, without expanding the number of questions greatly. Thus the prototype retains the usual weekly hours question.<sup>9</sup>

**EARNINGS FROM EMPLOYMENT.** To accurately measure earnings it is essential to phrase questions carefully (for example, referring explicitly to the last pay period rather than to "usual earnings") and to include the full range of questions needed to pin down the respondent's earnings both in cash and in kind. For cash earnings it is vital to ask explicit questions about respondents' income and payroll taxes, tips and gratuities, and bonuses. In addition, respondents must be

asked specific questions about any in-kind payments that they may have received in various forms. The draft employment modules are designed to collect a full set of data on respondents' jobs in the 7-day reference period but only summary data about cash wages and in-kind payments from respondents' jobs in the 12-month reference period. Less detail is justified when studying earnings in the 12-month reference period, both because the 12-month reference period data have seldom been used in detailed wage studies and because experience shows that respondents find it difficult to remember detailed earnings information that far in the past. Including more detail in the 12-month reference period section of the employment module also contributes to respondent fatigue, increasing the probability of inaccurate or missing information later in the interview.<sup>10</sup> Concern about respondent fatigue may also lead survey designers to decide to ask detailed questions about only one of the respondents' jobs in the 7-day reference period.

**ON-THE-JOB TRAINING.** The few questions about on-the-job training that were included in previous LSMS surveys were in the education module. They are better included in the employment module, however, because in this module reports of training can be linked with the employers that provide training and the occupations that require training, as well as with the individuals who receive training.

Measuring training is inherently difficult. Because much on-the-job training is very informal, recipients might respond "no" if they were simply asked whether they had ever received any training. Although asking this simple question is nevertheless useful for identifying formal training, less direct questions should also be asked that shed light on whether workers are in jobs in which their productivity increases as they watch others and practice their skills. The proposed module contains some questions for assessing the importance of informal training (associated with both getting better at doing a particular occupation and progressing to new occupations). However, it should be noted that these questions have not been field tested in developing countries and are in great need of pretesting. They are designed with an eye to avoiding the ambiguities in similar questions on informal training that were included in some U.S. surveys about which researchers have complained (Sicherman 1990).

**CONDITIONS OF EMPLOYMENT.** Given the multipurpose nature of LSMS surveys, the employment module restricts itself to qualitative measures (for example, whether the individual is entitled to a pension) of a few nonwage benefits and working conditions. Collecting quantitative data (such as size of pension and vesting requirements) is beyond the scope of the survey. With regard to job tenure, care should be taken to word questions so as to make clear that the desired response pertains to the amount of time the respondent has spent working for his or her current employer (in any occupation) and not the time he or she has spent doing the current sort of work. Finally, the draft module contains questions both about types of formal legal contracts (which must, of course, be modified to account for local employment contract law) and about informal arrangements between workers and employers.

#### *Facilitating Study of the Determinants of Labor Market Outcomes*

The first section indicated the potential usefulness for policy analysis of empirical studies of the determinants of wages and other labor market outcomes. This section briefly discusses how to facilitate such studies by collecting adequate information on potential determinants. While some of the determinants mentioned in the first section (such as gender, ethnicity, and geographic location) require little comment, two of them—the individual's schooling and the financial constraints the individual faced when entering the labor market—deserve further comment here. In addition, this section discusses several potential determinants of labor market outcomes that were not mentioned in the first section. These determinants—the individuals' labor market experience and innate ability, characteristics of the individuals' current households, and working conditions of the individuals' jobs—are seldom of direct policy interest, but are useful for getting good estimates of policy-relevant effects, either because they help avoid omitted variable bias in the estimation of policy-relevant effects or because they can serve as valuable instruments in attempts to control for endogeneity (see Chapter 26 on econometrics).

**SCHOOLING.** Quality of schooling received is often as important as years of schooling in explaining labor market outcomes. Simple measures associated with school quality include indicators of whether the last

school the individual attended was public or private, whether it was rural or urban, and perhaps what language was used in that school. It may also be possible to ask simple questions about the availability of textbooks and running water in the last school attended and whether the school was a completely enclosed structure. All of this information can be collected in the education module of a multitopic household survey.

**EARLY CAREER FINANCIAL CONSTRAINTS.** As indicated in the first section of this chapter, collecting indicators of the financial constraints facing the household in which the individual lived when he or she entered the labor market would give analysts the chance to assess the theory (which has not been studied very much) that poverty tracks workers into careers with little potential for advancement. In addition, collecting such data could be useful for a technical econometric reason; such measures would provide analysts with valuable and rare "instruments" to use in some econometric procedures for dealing with simultaneity biases (see Chapter 26 on econometrics), because in some contexts financial constraints variables could be expected to influence occupational or sectoral choice but not affect current wages directly.

In an attempt to assess the severity of financial constraints in the family that launched the individual into the labor market, it might be useful to modify the parental background section of the household roster to incorporate several indicators of the financial circumstances in which the individual found him or herself at the age of 15. (In some countries a younger age may be appropriate.) Assuming that the questionnaire already contains questions on parental education, (an important determinant of household earnings), it would be possible to control better for per capita disposable income by adding a question sequence such as: "Now I would like to ask you a few questions about when you were 15 years old. Was your father alive? Was your mother alive? How many brothers and sisters were living with you?" Other potential questions include questions about whether at that time household members were in good health, what land holdings the family had, the size of the family's house, and other simple indicators of wealth. Such questions have not been included in previous LSMS surveys, but experimenting with such questions would be valuable.

**LABOR MARKET EXPERIENCE.** Individuals' labor market experience is thought to be an important determinant

of wages and other labor market outcomes, because it reflects the opportunities individuals have had to acquire skills informally on the job, to increase wages by moving between jobs, and to improve their labor market standing in other ways. Failing to include good measures of experience in studies of wage determination can lead to misleading inferences about the effects of other characteristics on wages, with potentially serious consequences for policy analysis. For example, in countries where schooling attainment has been rising over time, the simple empirical relationship between wages and schooling in cross-sectional data, which assesses how much higher wages are for individuals with more schooling, would tend to understate the effect of schooling on wages. This is because the individuals who have more schooling also tend to be younger and have less experience (see discussion of omitted variables bias in Chapter 26). A regression that controls for experience as well as schooling would give a more accurate picture of the schooling effect.

Researchers studying the determinants of labor market outcomes who lack good measures of individuals' labor market experience often construct a measure of "potential labor market experience," which is equal to the individual's age in years minus the years he or she spent in school (including years repeated) minus the five or six years before entering school. This is an accurate measure only for individuals who have worked every year since leaving school and is probably much less accurate for women than for men. The draft employment modules allow for collection of more accurate experience measures.

**INNATE INTELLIGENCE.** Innate intelligence can influence wages and occupational choices for obvious reasons. Not being subject to policymakers' influence, the effects of innate intelligence on labor market outcomes are not of direct interest to policymakers. Again, however, failure to include measures of innate intelligence may lead to misleading policy inferences. For example, if individuals with higher innate intelligence tend to get more schooling, simple empirical relationships between wages and levels of schooling, which do not control for differences in innate intelligence, might overstate the wage effect of schooling.

Measuring workers' innate physical, mental, and social abilities is inherently costly, and even the best measures fall far short of measuring all of the traits valuable to employers. Research from developed and

developing countries seems to suggest that controlling for the abilities captured by feasible ability measures is unlikely to change inferences on many wage regression coefficients (Strauss and Thomas 1995). Thus, while it might be interesting to collect ability measures occasionally, collecting them in all surveys is probably not warranted.

**CHARACTERISTICS OF THE INDIVIDUAL'S CURRENT HOUSEHOLD.** Information on the structure of the individual's current household (for example, the numbers of dependents and of male and female workers) and on nonlabor income may be of great use in helping researchers deal with two econometric problems that plague studies of wage determinants: endogeneity of wage equation regressors such as indicators of sector or employer type and endogeneity of selection into the wage-earning subsamples on which wage regressions must be run (see Chapter 26 on econometrics and Schaffner 1997). In dealing with both problems it is crucial to find variables that might affect whether or not the workers are wage earners, the sector in which they work, and the employer type for which they work, but that do not affect wages directly. Household structure variables are likely to satisfy such requirements. They are unlikely to be known by employers or taken into account during wage setting, but, by influencing the value of potential workers' time in the home, they are likely to affect labor supply and sector choice decisions. Variables describing household structure may be derived from household roster information. For discussion of the measurement of nonlabor income see Chapter 11.

**WORKING CONDITIONS.** Some employer, workplace, and job characteristics are not themselves considered lines along which segmentation takes place but are potentially important controls in studying segmentation. For example, wages for comparable workers may be higher in an industry to compensate workers for working conditions that are less attractive in some way. Without controlling for working conditions, analysts might incorrectly conclude that there is segmentation along industry lines. Thus having data on working conditions (as well as on worker productivity factors such as schooling, training, and experience) to include in the wage regression allows analysts to make more accurate inferences about the potential severity of segmentation (Schaffner 1998).

## The Employment Module

This section introduces three potential employment modules: a short module, a standard module, and an expanded module with a variety of possible extensions. When survey designers are planning a single survey, they should not include all extensions simultaneously, as that would make the questionnaire too long. All questions should be pretested to ensure that they are meaningful for a wide range of workers (agricultural, rural nonagricultural, urban self-employed, and urban employed in both small and large establishments) in the country studied. The words and phrases in *italics* must be replaced by (lowercase) phrases relevant in a particular country before pretesting.

### *Links to Research Questions*

Table 9.1 links each issue in the first two sections of Box 9.1 with the relevant questions in the prototype modules and indicates how well the issues can be addressed using data from short, standard, and extended modules. The objectives listed in the first column of Table 9.1 are abbreviated references to the objectives listed in the first two sections of Box 9.1. Question numbers indicated under “Data needs” refer to questions in the “All extensions” version of the module. Separate notations are made under the standard module heading only when the standard module contains additional questions beyond those contained in the short module. The standard module column lists only the questions available in the standard version but not the short version. The table contains notations under the extensions column only when the extensions improve on the attainment of a research objective, and lists in this column only the additional questions available in the expanded module.

### *Age Limit*

The employment module should be administered to all household members above a specified age. Age limits in past LSMS surveys have ranged from 6 to 14. Survey designers should set this limit in accordance with local circumstances. Setting the limit at 10 or 12 years of age may be appropriate in many countries. Survey designers might occasionally want to lower the limit to 6 to study child labor,<sup>11</sup> in which case they might also want to include questions that check the respondents’ age and

steer younger workers into a shorter sequence of questions more relevant to them.

### *Respondent*

As in previous LSMS surveys, interviewers should attempt to obtain information about each individual directly from the individual himself or herself. After, say, two unsuccessful attempts to do this, interviewers may obtain information from a parent or another household member. Analysis of several previous LSMS surveys suggests that such “proxy respondents” are likely to be less accurate and less complete than the individuals themselves. Proxies appear less likely to report the individual as working or as having a second job and less likely to supply complete wage information. Among people who are reported as working, however, reported hours of work are higher for individuals whose hours are reported by proxies than for individuals who report their own hours. This is consistent with the hypothesis that individuals whose information is reported by proxy work more than other individuals (perhaps because their work keeps them away from home when the interviewers call) but that proxies are more reluctant to report the individuals as working, either because they do not want to answer a large number of questions about someone else (after possibly already answering all the same questions about themselves) or because they are uninformed or have a different perception about what constitutes the individual’s work.

### *Time Use Questions*

Household members make contributions to household welfare not only by working for wages or in self-employment, but also by engaging in unremunerated child care, cooking, cleaning, and other household production activities. Understanding household demands for such activities may be especially important for understanding household choices about education and work activities of children and about women’s employment and self-employment activities. Chapter 22 discusses the merits of collecting data on time used in such activities within a broader time use module. When such a module is not included in the survey, the employment module is a natural place in which to include several short questions on housework. Indeed, many previous LSMS employment modules have included such

**Table 9.1 Summary of Data Requirements by Research Objective**

Research objective	Short version		Standard version		Extensions (expanded version)		Other requirements
	Prospects for analysis	Data needs	Prospects for analysis	Data needs	Prospects for analysis	Data needs	
1. Labor force participation, employment, unemployment past 7 days; employment in past 12 months	Excellent	A2–A12					
2. Participation in various sectors of the economy	Excellent	B1–B3, C2–C7, C14, C24–C29, C34, D2–D7, D12, D21–D26, D31	Excellent	B3, B6			
3. Economic activities of poor	Excellent	Same as for 1 and 2					Consumption module measures of poverty
4. Hours of work	Good	B7–B10					
5. Unemployment duration and search methods	None		Good	A18–A22	Good	F21–F24	May need relatively large sample for accuracy.
6. Underemployment and on-the-job search methods	None		Good	A14–A17			
7. Average hourly earnings in wage employment	Good	C12–C21, C32–C41, D10–D18, D29–D36	Excellent for 7-day reference period; good for 12-month reference period	C10–C37, C58–C74			
8. Average hourly earnings in self-employment							Must be done in Household enterprise and Agriculture modules.
9. Annual earnings in wage employment	Good	Same as 7 plus B7–B10, C43–C44, D38–D39	Excellent				
10. Annual earnings in self-employment							Must be done in Household enterprise and Agriculture modules.
11. Participation in formal training	None		Good	C13, C48, C49	Good	C69–C73	
12. Participation in informal training	None		None		Good	C66–C68	
13. Linking training with employer and worker types	None		Good for formal training	Same as for 2 and 12	Good	C61–C65	Roster; Education
14. Nonwage benefits and nonwage job features	None		Good	C38–C40, C75–C77	Good	C19–C21	

(Table continues on next page.)

**Table 9.1 Summary of Data Requirements by Research Objective (continued)**

Research objective	Short version		Standard version		Extensions (expanded version)		Other requirements
	Prospects for analysis	Data needs	Prospects for analysis	Data needs	Prospects for analysis	Data needs	
15. Job length and nature of turnover	None		Excellent	C41–C42, C78, D15–D16, D20–D23	Excellent	F40–F42	
16. Participation in various contractual relationships	None		None		Excellent	C6, C7, C15–C18, C34, C80, C85–C87, D6, D7, D10–D13, D38–39, F11–F13, F35–F36	
17. Labor market effects of schooling, health, nutrition	Good	Same as for 1, 2, 4, 7, 9	Excellent	Same as for 2, 5, 7, 11, 14, 15	Excellent	Same as for 5, 11, 12, 14, 15, 16	Roster; Education, Health, Nutrition
18. Labor market effects of training and experience	None		Good for formal for 11 training	Same as	Excellent	A25–A29, C57–C65	Roster; Education
19. Labor market implications of contractual relations	None		None		Excellent	Same as for 16 and 17	Roster; Education
20. Discrimination	Fair	Same as for 1, 2, 4, 7, 9	Good	Same as for 2, 5, 7, 11, 14, 15	Excellent	Same as for 5, 11, 12, 14, 15, 16	Roster; Education
21. Intergenerational mobility	Fair	Same as for 2.					Roster; parental occupation
22. Labor market effects of financial constraints	Fair	Same as for 1, 2, 4, 7, 9	Good	Same as for 2, 5, 7, 11, 14, 15	Excellent	Same as for 5, 11, 12, 14, 15, 16	Roster; Education, plus expanded parental background (including proxies for financial constraints at age 15).
23. Labor market segmentation	Fair	Same as for 1, 2, 4, 7, 9	Good	Same as for 2, 5, 7, 11, 14, 15	Excellent	Same as for 5, 11, 12, 14, 15, 16, plus C19–C21	Roster; Education, Parental variables .
24. Labor market effects of community differences	Good	Same as for 1, 2, 4, 7, 9	Excellent	Same as for 2, 5, 7, 11, 14, 15	Excellent	Same as for 5, 11, 12, 14, 15, 16	Roster; Education, plus community module; better with panel or with repeated cross-section using same communities.
25. Change over time in labor force status and sector	Excellent	Same as for 1, 2			Possible without repeating cross-section or panel	F2–F12	Retrospective, repeated cross-section or panel data

**Table 9.1 Summary of Data Requirements by Research Objective (continued)**

Research objective	Short version		Standard version		Extensions (expanded version)		Other requirements
	Prospects for analysis	Data needs	Prospects for analysis	Data needs	Prospects for analysis	Data needs	
26. Change over time in average hourly earnings	Good	Same as for 7	Excellent for 7-day reference period jobs, good for 12-month reference period jobs	Same as for 7			Repeated cross-section or panel data
27. Change over time in nonwage job features	None		Good	Same as for 14	Good even without repeated cross-section or panel	F14–F16	Retrospective, repeated cross-section, or panel data
28. Change over time in contractual relations	None		Good	Same as for 16	Good even without repeated cross-section or panel	F11–F13	Retrospective, repeated cross-section, or panel data
29. Wage effects on labor supply and sector choice	Fair	Same as for 1, 2, 4, 7	Good	Same as for 2, 7			Repeated cross-section or panel with sufficient wage variation
30. Occupational and sectoral mobility	Good	Same as for 1, 2			Good even without panel.	F2–F6	Retrospective or panel
31. Relative success of job search methods.	None		Good	Same as 1, 2, 5, 6, 7			Panel
32. Income mobility	Good	Same as 9	Good	Same as 9			Panel, as well as complete data on income not from wage employment
33. Winners and losers associated with policy change	Good	Same as 1, 2, 7, 9	Good				Panel, as well as complete data on income not from wage employment
34. Career path effects of training and early good jobs	None		Good	Same as 1, 2, 7, 9, 11	Good	Same as 11, 12, plus F17	Retrospective or panel data
35. Labor market effects of risks and shocks	Good	Same as 1, 2, 7, 9					Panel of sufficient length; good agriculture and health modules for characterizing shocks
36. Relative success of job search by labor force status and location of residence	None		Good	Same as 1, 2, 5, 6, 7			Panel that follows migrants

Source: Author's summary

questions. Housework questions could be added after Question A12 in the short version of the module, and in comparable locations in the other versions. They might read: “During the past 7 days, have you done any work around the house such as cook-

ing, cleaning, washing clothes, fetching water or firewood, or performing maintenance work on your home?” (If the answer is no, the next question is skipped.) “During the past 7 days, how many hours did you do this kind of work?” “During the past 7

days, have you at times been the main person caring for or watching over one or more young children?” (If the answer is no, the next question is skipped.) “During the past 7 days, for how many hours were you the main person caring for or watching over one or more young children? DO NOT COUNT TIME WHEN CHILDREN ARE SLEEPING.”

### Question Counts

As indicated by the skip codes in the prototype questionnaire, many questions are relevant only for certain subsets of the sample. Table 9.2 provides rough calculations of the average number of questions to which individuals of various kinds should respond based on LSMS survey data for Peru and Cote d’Ivoire in 1985–86. Details of the assumptions on which this table is based are available from the author. The most important assumptions are that:

- 60 percent of individuals in the age range worked in the previous 7 days.
- 20 percent of those who did not work in the previous 7 days had a permanent job.
- 80 percent of those who worked in the previous 7 days—and 10 percent of those who did not—worked five years ago.
- 70 percent of jobs are in self-employment.
- 20 percent of all workers had some work in the previous 12 months that they were not performing in the past 7 days.

The question counts include checking questions, which the interviewer must answer but which are not read aloud to respondents. The first column averages over all individuals, whether they are employed by others, self-employed, unemployed, or out of the labor force.

### Notes on Proposed Questionnaires

This section gives detailed instructions about administering each set of questions included in the prototype employment module. It begins with the short version of the module and then discusses the additional questions asked in standard and expanded versions.

#### Short Version

**PART A. LABOR FORCE PARTICIPATION.** In response to this set of questions, interviewers fill out one line for each household member over the designated age. The

identification code in the first column identifies the individual by the number assigned to him or her in the household roster.

**A1.** This question indicates which responses are supplied by proxies rather than by the individuals themselves. Keeping a record of proxy responses is useful not only for assessing the quality of the survey but also because in some econometric applications researchers may want to test whether their results are affected by the inclusion or exclusion of proxy data.

**A2–A7.** These questions determine whether workers were employed in the previous 7 days and during the previous 12 months. To prevent underreporting of employment, the interviewers give respondents detailed prompts that are likely to make them think about and report some activities that they might not ordinarily consider “work” or “employment.” As in past LSMS surveys, this draft module uses the International Labor Organization’s definition of employment, but to improve accuracy the questions here have been amplified. They could be amplified even further. For example, the phrase “for at least one hour” could be inserted into each question to prevent the underreporting of activities that are only carried out part-time. Questions A2 and A3 could be amplified by adding a sentence such as: “This work might have been for payment in cash or in kind, in exchange for labor or a reduction of debt, or as part of an apprenticeship or on-the-job training program.” Even if such a sentence is not included, interviewers should be trained to recognize this wide variety of activities as “work.”

Questions A6 and A7 should contain as many phrases as is necessary to cover the wide range of potential self-employment activities. The draft questionnaires use just two: “work on own account” and “work in a household business enterprise.” Both here and in the household enterprises module, careful preliminary research should be done to determine which phrases best capture all self-employment activities.

Because standard definitions of employment use a reference period of the previous 7 days, this is the reference period used in the draft module. An alternative would be to refer to “last week, from Sunday through Saturday.” However, this leads to longer recall periods for respondents interviewed later in the week. Questions A3, A5, and A7 are abbreviated relative to A2, A4, and A6, but the same comments apply.

**Table 9.2 Estimated Question Counts per Individual**

Version	Average over all individuals	Individuals with one job in wage employment and the same job five years ago	Individuals with one job in self-employment and the same job five years ago	Individuals with one job in wage employment and one in self-employment, the same jobs five years ago	Individuals with one job in wage employment and a different wage employment job five years ago
Short	30.7	40.2	26.0	54.2	40.2
Standard	39.7	62.0	31.7	77.3	62.0
Standard plus Extension 1 (how found current job, A24)	39.8	62.0	31.7	77.3	62.0
Standard plus Extension 2 (experience, A25–A29, C57–C60)	43.4	67.4	34.7	82.7	67.4
Standard plus Extension 3 (five-year retrospective, A30–A31, E1–E5, F1–F42)	51.7	72.3	42.0	94.9	88.9
Standard plus Extension 4 (contract and relationship with employer; C6–C7, C15–C18, C80, C85–C86, D6–D7, D10–D13, D38–D39)	41.5	68.3	31.7	83.6	68.3
Standard plus Extension 5 (working conditions, C19–C21, C87)	40.1	64.0	31.7	79.3	64.0
Standard plus Extension 6 (skill acquisition, C61–C65, C113–C115)	40.3	64.8	31.7	80.1	64.8
Standard plus Extension 7 (informal training, C66–C68)	40.1	64.0	31.7	79.3	64.0
Standard plus Extension 8 (length of formal training, C70, C71, C73)	39.7	62.6	31.7	77.9	62.6

Source: Author's calculations from Côte d'Ivoire and Peru survey data.

**A8.** This question asks the interviewer to identify which respondents did and did not work during the previous 7 days, so that each group can be asked appropriate questions.

**A9.** This question asks individuals who did not work during the previous 7 days whether they have a permanent job. Individuals who have a permanent job but were absent from work during the previous 7 days should be counted as employed according to the International Labor Organization's definition.

**A10.** This question asks individuals who were absent from a permanent job why they were absent. This helps analysts interpret the report that a respondent has permanent job but worked no hours during the previous week. It also serves as a check on that report, because the answer may reveal that the respondent misunderstood the previous question.

**A11.** This question is necessary for identifying individuals who are in open unemployment by standard definitions. Sometimes survey designers replace the word "work" with "paid work," but this may cause

respondents to focus too narrowly on wage employment. The question could be amplified by including the phrase "whether by consulting newspapers, employment agencies, employers, friends, or relatives."

**A12.** This question asks individuals who did not work and did not search for work why they did not search for work (that is, why they were not in the labor force). Qualitative research and pretesting should be used to refine the list of potential responses.

**A13.** This question is used to determine what the interviewer should do next. If the respondent has worked in the past 12 months, the interviewer should proceed to Part B and continue interviewing that person. If the respondent has not worked in the past 12 months, no more employment questions need to be asked of that person.

**PART B: OVERVIEW OF WORK IN THE PREVIOUS 7 DAYS AND THE PREVIOUS 12 MONTHS.** In response to the questions of this section, interviewers fill out one line for each work activity reported by each individual who was identified by Question A8 as having worked or had

**Box 9.2 Cautionary Advice**

- *How much of the suggested module is new and unproven?* Most of the questions in the short and standard modules are similar to those used in previous LSMS surveys and in many labor force surveys around the world. Some questions have been modified slightly to increase specificity and clarify distinctions between occupations and jobs. The questions related to experience, contracts, relationships with employers, skill acquisition, and training are new and largely unproven. The draft modules are organized somewhat differently from previous LSMS employment modules. Previous LSMS employment modules have allowed respondents to identify their main and secondary work in the past 7 days and past 12 months, and asked detailed questions about each of these jobs immediately after they were identified—so that secondary jobs and 12-month reference period jobs were identified only after workers had already been asked many detailed questions about the main job of the last 7 days. The current modules contain brief overview sections on all work (not just main and secondary jobs) done in the past 7 days and in the past 12 months before asking detailed questions about main and secondary jobs in the past 7 days and 12 months; in addition, the current modules require interviewers, rather than respondents, to identify main and secondary activities in each reference period. The only potential difficulty that this introduces is that interviewers must calculate the number of hours worked by household members in each occupation in a given reference period in order to know which two occupations were most important during that reference period. They must then refer the respondent to his or her previous answers to questions about occupation and industry. In light of these challenges, it is important to train interviewers carefully.
- *How well has the module worked in the past? As far as we know, the employment module has worked reasonably well in the past.* Validation studies mentioned in this chapter suggest, however, that the greater specificity in the phrasing of questions in the draft modules should be useful for improving accuracy. Also, asking respondents to give brief overviews of work activities before asking them to answer many more detailed questions should lead to more complete and accurate description of work activities.
- *What parts of the module most need to be customized?* Phrases rendered in italics in the prototype modules are especially in need of customization. In addition, qualitative research would be useful for determining the in-kind payments, nonwage benefits, and working conditions about which specific questions should be asked.

a job during the previous 7 days, or by Question A13 as having worked in the last 12 months. Lines are also filled out for any work done in the past 12 months that was not done in the previous 7 days. The interviewer must fill in the identification code of the individual and must use the “activity number” column to number sequentially the activities reported by the respondent.

Part B briefly identifies all the kinds of work (occupations) in which respondents have been engaged during the previous 7 days. The interviewer determines which two of these occupations were the most important (in the sense of having the most hours devoted to them). Part C then asks more detailed questions about the respondent’s main job (that is, work done under the employer or household enterprise for which the respondent has worked the most hours) in each of the two most important occupations during the previous 7 days. Part B has been kept quite short in the hope that this will increase the probability of respondents giving a complete account of the work they did during the previous 7 days. If, however, survey designers plan on asking only a few more detailed questions about 7-day reference period jobs, and if they are willing to ask those questions about all work activities in the reference period (rather than just the two most important ones), the questions could be added to this section and Part C could be eliminated.

Part B also determines which occupations were the most important in the past 12 months. If they are the same as the most important occupations in the past 7 days, the information on them is gathered in Part C. If data are not collected on either of those jobs in Part C, the information is collected in Part D.

Respondents are asked to report on any permanent jobs they may have held during the previous 7 days even when they actually did no work during that time. It is important to include these questions in the section dealing with the 7-day reference period rather than the section dealing with the 12-month reference period because only in the 7-day section can detailed information about these jobs can be obtained. It seems likely that jobs from which workers might be temporarily absent are systematically different from other jobs. Paid vacations, for example, are likely to be associated with relatively high-paying, formal jobs. Analysis of wages and other labor market outcomes would be incomplete without detailed information on such jobs.

**B1–B2.** These questions identify the respondent's occupation and industry. Interviewers write out descriptions; the responses are coded later.

**B3.** The main purpose of this question is to identify nonfarm self-employment activities so that the respondents involved in such activities can be asked to match the activities to the household enterprises described in the household enterprise module.

**B4.** This question links data on self-employment in the employment module to data on self-employment in the household enterprise module. This is important both for linking the data in the two modules that pertain to the same activities and (perhaps more importantly) for checking that detailed data are being collected in the household enterprise module for all activities that are considered to be “self-employment” activities in the employment module (because the employment module elicits very little information about self-employment activities).

**B5–B6.** Question B5 helps the researcher verify any cases in which the respondents claim that the number of hours that they worked overall in an activity exceeded the number of hours that they worked for their main employer in this activity during the previous 12 months (which is only possible if they have more than one employer in this activity). B6 does the same for the past 7 days. These questions are also useful for identifying casual workers who work for several employers during one 7-day period.

**B7–B10.** These questions describe the amount of time the respondent spent working in this occupation during the previous 7 days and the previous 12 months. This enables a rough estimate of the number of hours that the respondent worked in a variety of reference periods. Weekly and annual earnings in this occupation can be estimated by multiplying the respondent's average hourly earnings from his main employer in this occupation by these total labor supply numbers. Interviewers should be trained to prompt respondents who find these questions hard to answer and to help respondents add up hours or weeks.

**B11.** This question asks interviewers to calculate numbers to which they will refer in Question B16 when they ask respondents about their main and secondary

jobs during the previous 12 months. If survey designers decide to ask detailed questions about all of the respondents' jobs rather than just their main and secondary jobs, this question is unnecessary. There would also be no need for this question if this objective way of picking the two most important jobs were replaced by a more subjective method—for example, if after describing all their work activities, respondents were simply asked which two were the “most important.”

**B13.** This question determines whether another row in this table pertaining to 7-day reference period activities should be completed for a particular respondent. If not, the interviewer should move on to B14.

**B14.** This question asks the interviewer to flag the two activities to which the respondent devoted the most hours during the previous 7 days. The interviewer will later refer to this column when determining which activities should be the subject of Part C.

**B15.** This question determines whether another row in this table pertaining to 12-month reference period activities should be completed for a particular respondent. If not, the interviewer should move on to B16.

**B16.** This question asks the interviewer to flag the two activities to which the respondent devoted the most hours in the last 12 months. The interviewer will later refer to this column when determining which activities should be the subject of Part D. Note that this question requires the interviewer to consider all activities the individual reports in this section, whether relevant to the 7-day reference period or only to the 12-month reference period.

**B17–B22.** In answering these questions, the interviewer is determining which sections of Parts C and D must be administered to this respondent and is recording the results in one place. The answers to these questions appear only on the last line of Part B for each respondent. The interviewer will refer to these answers later when determining whether to administer various sections of Parts C and D to each respondent. When questionnaires are computerized, these answers can be filled in automatically. When they are not, interviewers will require training to understand the intent and to be able to complete them quickly.

**PART C: MAIN AND SECONDARY JOBS IN THE PREVIOUS 7 DAYS.** Interviewers should fill out one line for each person who has worked (or had a job that he or she was temporarily away from) during the past 7 days. Part C collects information on the activities marked in B14 with a 1 or a 2.

Questions about the 7-day reference period have been split into two sections: the overview in Part B and the detailed questions of Part C. There are three reasons for this split. First, the overview elicits information about all work activities of the past 7 days, while the detailed questions are asked about only two jobs. Second, it is useful to keep the overview section short so that a complete accounting of activities may be recorded before respondents become weary of a long list of questions. Finally, the focus in the overview is on occupations, while the focus in the detailed questions is on the main job within each occupation. Similar points can be made about splitting questions about the 12-month reference period in Parts B and D and splitting the retrospective questions into Parts E and F.

**C1.** This question is used to establish which occupation mentioned in Part B will be discussed in the questions that follow. (This occupation is specifically mentioned in C2 to remind the respondent.)

**C2.** This question does three things. First, it asks the respondent to return to thinking about one specific activity that he or she performed during the previous 7 days. Second, it shifts the respondent's focus from an occupation or activity (which might have been performed for more than one employer) to a job (for a specific employer or household enterprise). Third, it asks where this work was carried out, which some analysts and policymakers take as an indicator of the degree of formality of the job.

**C3.** The main purpose of this question is to differentiate between jobs for which the rest of the employment module is relevant and jobs for which it is not. The skip codes here indicate that the rest of the module will refer either to work that the respondent does for an employer who is not a member of the same household as the respondent or to work that the respondent does for a household enterprise for which he receives a wage (presumably from another household member in charge of the enterprise). If survey designers prefer not to put the rest of the module's

questions to paid workers in household enterprises, the skip codes should be adjusted accordingly.

**C4.** The answer to this question should indicate the respondent's broad sector of employment. Where relevant, a large rural public works program (such as the Employment Guarantee Scheme in India) should be designated by name.

**C5.** This is a standard question about employer size. In some countries employer size is an important legal distinction, with certain regulations applying only to employers with at least some minimum number of workers. Even where there is no legal exemption of small firms from regulations, size is often taken as an important indicator of likely compliance with regulation. In many past LSMS surveys this question was addressed only to workers employed in the private sector. Such a limitation could be imposed by adding an appropriate skip code to C4. No such code is included in the draft module because information on size may be just as important for studying wage determination in the public sector as it is in the private sector. Rather than asking for an exact number, the question could provide several size categories from which the respondent can choose.

**C6.** This is a standard question used to distinguish jobs that are covered by labor law from those that are not covered.

**C7.** This question will yield answers that are more useful for studying the effects of unions on wages than the question "Are you a member of a union?" would be, because unions often strike wage deals that apply not only to unionized workers but also to nonunion workers in the same firm or industry.

**C8–C9.** These questions measure the distance between a worker's residence and place of work. This is useful in studies of wage determination because employers who have to recruit workers from a wide area may have to pay high wages to attract these workers (who must incur high commuting costs). These questions may also be useful for distinguishing temporary migrants from commuters.

**C10–C11.** These questions determine whether the extensive set of questions on wages should be admin-

istered and elicit information on why some workers report receiving no wages. Workers who receive no payment should skip all questions about cash and in-kind payments and about nonwage benefits. If unpaid workers are not of direct interest to survey designers, the questions on training, experience, and skill acquisition can also be excluded for these workers by skipping to C18 rather than to C15.

**C12.** This is the main question about cash earnings. The accuracy of this measure is greatly increased by referring to the last payment (or the first one if the respondent has yet to receive one) rather than to the “usual” payment.

**C13.** This question is a simple attempt to increase the accuracy of the calculation of average hourly earnings beyond what was attained in previous LSMS surveys.

**C14.** This question provides an additional indicator of a job’s formality by revealing whether the pay is subject to income or social security tax. Previous LSMS surveys were rather vague about whether they measured before-tax or after-tax pay.

**C15–C16.** These questions elicit information on any other payments—for example, in the form of meals, housing, and clothing. Qualitative research should be done to determine the most important types of in-kind payments in the region surveyed.

**C17–C21.** These questions elicit data on hours worked for this employer during the previous 7 days and 12 months, if the respondent worked in this occupation for more than one employer in the past 12 months. C18–C21 differ from B7–B10 in that they refer to work for a particular employer rather than work for any employer in a particular occupation. Analysts need a figure for the number of hours that the respondent has worked for this employer in order to measure the respondent’s average hourly earnings from this employer. They need a figure for the number of hours that the respondent has worked in the occupation to extrapolate the respondent’s earnings from this employer to his or her total earnings in the occupation.

**C22–C23.** These questions determine whether the individual had a second occupation in the past 7 days.

**C24–C41.** These questions are asked for all respondents with a second occupation in the past 7 days. These questions are a subset of Questions C2–C21; thus the corresponding comments above apply to them.

**C42–C44.** Any income from a third, fourth, or higher number of occupations in the past 7 days is reported here.

**C45–C47.** These questions determine whether the respondent should be asked questions in Part D and, if so, which question in Part D to begin with.

**PART D: MAIN AND SECONDARY JOBS IN PREVIOUS 12 MONTHS (WHEN DIFFERENT FROM JOBS IN PREVIOUS 7 DAYS).** In this section interviewers should fill out one line for any activity marked with a 1 or a 2 in Question B16 that has not already been discussed in Part C. These activities pertain either to year-round jobs that the respondent has left in the last 12 months or to jobs performed only seasonally. This section of the module asks for much less detail about the respondent’s earnings than does the 7-day reference period section, although it still attempts to gather the summary measures necessary for estimating the respondent’s annual earnings from wage employment. The comments on almost all of the individual questions in this section are the same as those that apply to the corresponding questions in Part C. See the introductory notes to Part C for a discussion of why 12-month reference period questions are separated into two sections (the overview in Part B and the detailed questions in Part D).

#### *Standard Version*

Many of the questions in the standard version are also in the short version, and virtually all of the questions in the short version are also in the standard version. For comments on questions that are in both versions, see the comments given above in the subsection that describes the short version. Comments are provided in this subsection for questions in the standard version that are not in the short version.

#### **PART A: LABOR FORCE PARTICIPATION.**

**A1–A2.** These questions indicate which responses are supplied by proxies rather than by the individuals

themselves. They also identify the proxy (Question A2). Keeping a record of proxy responses is useful not only for assessing the quality of the survey but also because, in some econometric applications, researchers may want to test whether their results are affected by the inclusion or exclusion of proxy data. Keeping such a record is also useful for studying whether and how proxy responses vary depending on the gender, age, and education of the proxy respondent. This can help determine how proxies should be chosen in future surveys. (See Chapter 4 on metadata.)

**A14–A15.** These questions ask people who did work whether they also looked for additional or replacement jobs. Affirmative responses to such questions are used by some analysts and governments as indicators of “underemployment.”

**A16.** This question checks whether the respondent conducted any on-the-job search, in which case A17 is relevant.

**A17.** This question asks what methods the respondent used for his or her on-the-job search. Survey designers may want to differentiate between public and private employment agencies. Intermediaries, who are often known by slang names, may be especially important in rural areas for finding work both locally and far away.

**A18–A22.** These questions ask about the duration of the respondent’s job search and the search methods used during any period of unemployment experienced in the previous 12 months. In earlier LSMS Surveys questions about the duration of unemployment and the search methods used were asked only of respondents who were unemployed during the previous 7 days—which tends to be a small sample. This draft module seeks to identify a larger sample by asking about any unemployment experienced during the previous 12 months.

**A23.** This question is used to determine what the interviewer should do next. If the respondent has worked in the past 7 days or 12 months, the interviewer should proceed to Part B and continue interviewing that person. If the respondent has not worked in the past 7 days or 12 months, no more employment questions need to be asked of that person.

**PART B: OVERVIEW OF WORK IN THE PREVIOUS 7 DAYS AND THE PREVIOUS 12 MONTHS.** See the general comments on the purpose of Part B in the subsection above that describes the short version of the employment module.

**B3.** This question will be of greatest interest in regions where there are large agricultural wage labor markets. Norms of payments and other working conditions—and their evolution over time—typically differ by crop. Thus analysts who want to make sense of data on agricultural wage labor markets will want to distinguish such workers by crop.

**B6.** This question is relevant only in countries with social security programs to which the self-employed may contribute. Although the question may seem out of place in the overview section, this is the most convenient place to put it because the preceding questions have already identified the group to which the question should be asked.

**PART C: MAIN AND SECONDARY JOBS IN THE PREVIOUS 7 DAYS**

**C8–C12.** These questions perform at least three functions. First, they make it possible to calculate the value of subsidized transport provided by employers. Second, they make it possible to measure the distance between a worker’s residence and his or her place of work. (This is useful in studies of wage determination because employers who have to recruit workers from a wide area may have to pay high wages to attract those workers in light of the workers’ high commuting costs.) Finally, this set of questions is useful for distinguishing temporary migrants from commuters. Where seasonal migration flows are of special interest, questions can be added here that are relevant only to workers who commute sufficiently infrequently or over sufficiently great distances. For example, such questions might identify the municipality in which the work was done.

**C13–C15.** These questions elicit information about apprenticeship fees. Their main purpose is to increase the accuracy of estimates of income and expenditures. In some countries the questions may also be useful for studying apprenticeships themselves, but in most countries a typical LSMS sample will not be big

enough to do this. Qualitative research should be used to determine local colloquial expressions for apprenticeship arrangements.

**C20–C22.** These questions provide an additional indicator of a job's formality by showing whether the pay is subject to income or social security tax. They also increase the accuracy with which after-tax pay can be measured. Previous LSMS surveys were vague about whether they were measuring before-tax or after-tax pay. The answers to Questions C20–C22 make it possible to calculate take-home pay for all workers but allow calculation of pretax pay only for workers whose initial answers reveal pretax pay. (This reflects the assumption that some workers may find it difficult to report their pretax pay, which they never see.) When survey designers believe workers can generally report both before- and after-tax pay, a question should be added asking respondents who have provided their after-tax pay about their pretax pay.

**C23–C24.** These questions improve the measurement of cash wages.

**C25.** This question is of interest in determining average hourly earnings as indicators of the extent to which employers can provide direct pay incentives for hard work.

**C26–C28.** These questions improve the measurement of cash earnings. The periodicity question is asked somewhat differently here than in previous LSMS surveys, in the hopes of eliciting accurate data.

**C29–C37.** These questions elicit information on several payments made in kind, including payments in the form of meals, housing, clothing, and "other." Qualitative research should be done to determine the most important types of in-kind payments in the region to be surveyed.

**C38–C40.** These questions elicit indicators of nonwage benefits.

**C41.** This question measures the length of respondents' tenure with their employer. This information is useful for studying job stability as well as wage determination. Although the question is phrased in a more precise way than in previous LSMS surveys, it could be amplified

further. For example, it could be rephrased as "For how long have you worked continuously for this employer?" or "For how long have you worked for this employer with breaks of no more than three months?" This draft module does not use such questions because the first may be confusing and second is clumsy.

**C42.** This question identifies seasonal workers who have multiyear arrangements with employers. The question is of greatest interest in regions where policymakers are concerned with seasonal migrant groups. Answers to the question may help researchers confirm the validity of data from individuals who report having worked for their current employer for many years but for only a few months during the previous 12 months.

**C48–C49.** These questions identify whether the worker has completed formal job training or is still in a "formal training period" in his or her job. Again, this is useful for studying wage determination. Identifying the occupations or industries in which formal training programs are important might be useful for identifying sectors in which analysts might wish to conduct case studies.

**C50–C51.** These questions determine whether an individual had a second occupation in the past 7 days.

**C52–C83.** These questions are asked of all respondents with a second occupation in the last 7 days. The questions are a subset of Questions C2–C48; thus the corresponding comments above apply to them.

**C84–C86.** Any income from a third, fourth, or higher number of occupations in the past 7 days is collected here.

**C87–C89.** These questions determine whether the respondent should be asked questions in Part D, and if so, which question in Part D to begin with.

**PART D: MAIN AND SECONDARY JOBS IN PREVIOUS 12 MONTHS (WHEN DIFFERENT FROM JOBS IN PREVIOUS 7 DAYS).** In this section, interviewers should fill out one line for any activity marked with a 1 or a 2 in Question B18. The questions in Part D ask for much less detail about the respondent's earnings than does the 7-day reference period section, although Part D

still attempts to gather the summary measures necessary for estimating the respondent's annual earnings from wage employment. The comments on almost all of the individual questions in this section are the same as those that apply to the corresponding questions in Section C. See the introductory notes to Part C for a discussion of why 12-month reference period questions are separated into two sections (the overview in Part B and the detailed questions in Part D).

**D21–D22.** These questions aim to find out why the jobs that were important to the respondent during the previous 12 months were not carried out in the previous 7 days. One reason might be that the respondent changed from one nonseasonal job to another, in which case the data will indicate the relative importance of quits and fires and the relative importance of job separations that do and do not cause workers to become unemployed. Another reason may be that the job in question was a seasonal job, in which case the data will indicate whether the seasonal diversification of activities is important in the country or region studied.

#### *Expanded Version*

Many of the questions in the expanded version are also in the short and standard versions, and virtually all of the questions in the short and standard versions are also in the expanded version. For comments on questions in the expanded version that are also in the short or standard versions, see the comments above in the subsections that describe those versions. The comments in this subsection are for questions that appear only in the expanded version.

#### **PART A: LABOR FORCE PARTICIPATION.**

**A23–A24.** These questions ask about channels through which people found jobs that they held in the past 7 days. (Such questions could instead be asked in Part C, where they could be associated with particular employers.)

**A25–A29.** These questions are used to measure total work experience, which tends to be an important explanatory variable in studies of wage determination. Question A25 ensures that only individuals who reported having no work during the previous 12 months are asked if they have ever worked.

**A30–A31.** These questions determine whether or not individuals were working five years earlier. Such questions could be left until the beginning of the five-year retrospective section, but it is useful to note which individuals were working five years ago early in the interview, before the respondents become tired. Fatigued respondents may be reluctant to admit having worked five years ago, knowing that such an admission is likely to open them up to a new battery of questions.

**A32–A33.** These questions help the interviewer determine which sections of the module, if any, to next administer to the current respondent.

**PART B: OVERVIEW OF WORK IN THE PREVIOUS 7 DAYS AND THE PREVIOUS 12 MONTHS.** See the general comments on the purpose of Part B in the subsection above that describes the short version of the employment module. Since all of the questions in Part B of the expanded version are also in the standard version, see the comments in the subsection on the standard version (as well as the comments in the subsection on the short version).

**PART C: MAIN AND SECONDARY JOBS IN THE PREVIOUS 7 DAYS.**

**C6–C7.** These questions ask about aspects of the worker's relationship with his or her boss. This is useful in some studies of wage determination.

**C15.** This question is of interest primarily in countries where certain intermediaries are exempted from various aspects of labor law.

**C16–C17.** These questions provide a more detailed look at the relevance of labor law. Labor laws typically defines several types of employment contract, making it difficult to fire workers hired under certain types of contracts and restricting the conditions under which other types of contracts may be used. In some countries Question C16 would be phrased quite differently. In Brazil it would be phrased as: "Do you have a signed workers' card (carteira assinada)?"

**C18.** This question is a rough attempt to determine the respondent's perception of how long the job would last at the time he or she first took the job. This is of inter-

est for determining whether jobs outside the scope of job security legislation (whether legally or illegally) are more or less stable than jobs within the scope of that legislation. Qualitative research should be used to determine the most useful way of asking this question.

**C19–C21.** These questions yield simple indicators of working conditions, which are useful in studying wage determination.

**C57–C60.** These questions measure the respondent's experience in his or her current industry (that is, with any employer in the same line of business, whether the respondent's work responsibilities were the same or different) and in his or her current occupation (that is, doing the same kind of work, whether for employers in the same line of business or a different one). The distinction between the two could be made clearer by including explicit references to the industry and occupation that the respondent reported in Questions B1 and B2, but this is cumbersome. The more important of the two experience measures is the industry-specific experience measure. The effects of this measure on wages are of interest for the light they shed on the nature of training problems. If experience acquired while working for different employers in the same industry increases a worker's productivity and wages in the current job, there is reason for concern that poaching problems lead to underinvestment in training within the industry, and that policies facilitating industry-based training cooperatives might help mitigate this problem.

Both industry- and occupation-specific experience measures are also useful for studying broader questions about wage determination, because it is important to control as thoroughly as possible for the skills that workers bring to their jobs (for example, in studying possible segmentation or discrimination).

**C61–C63.** These questions aim to shed light on the overall importance of skilled labor, on the relative importance of formal education in providing workers with the skills they need, and on whether one kind of enterprise in effect trains the staff of other enterprises (for example, informal household enterprises and small establishments providing skilled labor to larger, more formal establishments).

**C64–C65.** These questions identify where workers first started gaining experience within their occupations.

Again, the questions aim to shed light on relationships among different types of enterprises. If workers tend to stay with the employers who first provided them with experience and training, there is reason to think that training produces "firm-specific" skills that can be developed only under long-term employment arrangements. This may mean that labor laws and macroeconomic policies that make long-term employment arrangements unattractive discourage training.

**C66–C68.** These questions identify whether a worker is still in a "training period" in his or her job (either informal or formal), and if not, how long ago that period was. This is useful for studying wage determination and for assessing the importance of on-the-job training in the economy.

**C69–C73.** These questions identify whether the worker is still in a "formal training period" in his or her job and if not, how long ago such a period was. Again, this is useful for studying wage determination. Identifying the occupations or industries in which formal training programs are important could be useful for identifying sectors in which analysts might wish to conduct case studies.

**C74–C75.** These questions determine whether an individual has had a second occupation in the past 7 days.

**C76–C117.** These questions are asked of all respondents who have had a second occupation in the past 7 days. These questions are a subset of Questions C2–C72, so the corresponding comments above apply to them.

**C118–C120.** Any income in the past 7 days from a third, fourth, or higher number of occupations is collected here.

**C121–C123.** These questions determine whether the respondent should be asked questions in Part D and, if so, which question in Part D to begin with.

**PART D: MAIN AND SECONDARY JOBS IN PREVIOUS 12 MONTHS (WHEN DIFFERENT FROM JOBS IN PREVIOUS 7 DAYS).** In this section interviewers should fill out one line for any activity marked in Question B18 with a 1

or a 2 that has not already been discussed in Part C. This section of the module asks for much less detail about the respondent's earnings than does the 7-day reference period section, although this section still attempts to gather the summary measures necessary to estimate the respondent's annual earnings from wage employment. The comments on almost all of the individual questions in this section are the same as those that apply to the corresponding questions in Part C. See the introductory notes for Part C for a discussion of why 12-month reference period questions are separated into two sections (the overview of Part B and the detailed questions in Part D).

#### **PART E: OVERVIEW OF WORK DONE FIVE YEARS**

**PREVIOUSLY.** In this section, interviewers should fill out one line for each activity reported by each respondent who was identified by Question A30 as having worked five years ago. The interviewer must fill in the respondent's identification code and number the activities sequentially. Parts E and F are much shorter than the sections on current jobs because they emphasize describing the sectors in which the respondents' work activities took place, rather than quantifying the respondents' labor supply or earnings (which respondents may have difficulty recalling accurately). The comments on many specific questions are the same as those for corresponding questions in previous sections.

**E5.** Whereas respondents' first and second most important work activities in the previous 7 days or the previous 12 months were determined by the number of hours devoted to each work activity, the respondents' first and second most important work activities 5 years ago are determined simply by asking them.

#### **PART F: MAIN AND SECONDARY JOBS FIVE YEARS**

**BEFORE THE INTERVIEW.** In this section interviewers fill out one line for each activity marked with a 1 or a 2 in Question E5. The comments on most individual questions are the same as those for corresponding questions in previous sections. See the introductory notes to Part C for a discussion of why questions about this reference period are split into an overview (Part E) and detailed questions (Part F).

**F4–F5, F29–F30.** These questions determine whether the respondents' first and second most important work activities of five years ago are the same as their first and

second most important work activities during the previous 7 days or 12 months. Such a determination is useful for two reasons. First, it provides a more accurate way to infer which respondents did and did not change jobs over the last five years than simply trying to match industry and occupation codes. Second, it allows an interviewer to skip some questions about aspects of the respondent's job that are unlikely to have changed.

**F17–F18.** These questions are asked even if the respondent continues to work for the same employer at the time of the survey, because they pertain to features of employment that may have changed over time.

**F18.** This question checks whether the job is still going on, so that questions about why the job ended will be asked only of respondents for whom such questions are relevant.

#### **Notes**

The author gratefully acknowledges research assistance provided by Meera Mehta and Jaana Remes, and comments from many people, including Jere Behrman, Richard Blundell, Catherine de Fontenay, Paul Glewwe, Margaret Grosh, Anjini Kochar, Tom MaCurdy, Alberto Martini, Andrew McKay, Anne Royalty, John Pencavel, Jo Van Biesbroeck, and Wim Vijverberg.

1. Answering such questions is only a first step, because establishing the correlation of a policy change with labor market changes does not prove that the policy change caused the labor market changes. Giving careful thought to the timing of the changes and to the likely importance of other potential explanations for the labor market changes is often helpful in determining whether or not the policy change was an important cause of the labor market change.

2. The need for a worker to sacrifice earnings early on to obtain higher earnings later may arise for several reasons. The initial wage may be low because the worker is also receiving free training that will enhance his or her productivity. Alternatively, employers may use the promise of higher wages later to motivate workers to work hard or to provide workers with reasons not to quit. Sacrificing early wages by spending more time in open unemployment may also help workers find higher-paying jobs later in life.

3. The effective wage lies below the actual wage to the extent that program transfers fall as earned income rises. Nonlabor income is increased by the size of the maximum transfer.

4. A potential drawback to this approach is that it does not permit collection of detailed information about the individual's

work in the same occupation for two different employers. Comparing the wages earned by the same individual in the same occupation for employers in both the public and the private sector—or for work in both large and small establishments—can be useful for assessing the likely importance of segmentation, because it enables the measurement of wage differences that cannot be the result of differences in workers' abilities. However, the number of individuals performing the same job for employers in different sectors is unlikely to be large in a LSMS sample, so this is unlikely to be a major loss.

5. If wages depended only on variables that also belong in labor supply and job choice relations directly, predicted wages would be linear functions of other variables on the right-hand side of the labor supply and job choice relations, and the regressions would suffer from perfect multicollinearity. For more on simultaneity and the use of instrumental variables see Chapter 26.

6. Because changed wages today may influence work and sector choices later in life, one might want to estimate “dynamic” labor supply models that allow for the full range of such effects. Only panel data allow estimation of the full set of parameters required for tax and transfer policy analysis under such circumstances. Useful subsets of those parameters (for example, allowing prediction of the effects of permanent but not transitory policy changes) can be estimated with repeated cross-sectional data, especially when good estimates of consumption are available (see MaCurdy 1985). At this point the full dynamic models requiring panel data are tractable only under very strong assumptions that make them unattractive for studying developing countries.

7. Rodgers, Brown, and Duncan (1993) find that measurement errors are more likely to be correlated over time for an individual, and thus are less likely to lead to spurious transitions, when observations are collected retrospectively in one interview than when they are collected in interviews at different dates.

8. Panel data are sometimes thought useful for a very different reason. When economic conditions are not changing much over time, panel data allow analysts to use “fixed effects” econometric procedures, which under certain strong assumptions eliminate potential biases in the cross-sectional econometric relationships that arise out of the failure to measure all relevant individual characteristics. Such methods are less useful than is sometimes thought. They aggravate biases associated with measurement error (which are substantial), and they are based on assumptions (for example, that the unobserved characteristics do not change over time and that the relationship being estimated is stable over time) that are often faulty. Thus the more pressing rationale for collecting panel data is interest in intrinsically dynamic questions like those raised by issues 30–36 of Table 1. See Chapter 23 for a discussion of panel data uses.

9. International Labor Organization studies also suggest that accurate measurement of labor supply in a reference period is facil-

itated by the use of question sequences that require respondents to make a full accounting of their time use in the reference period (Hussmans 1990). Unfortunately this, too, is beyond the scope of the employment module of a living standards survey. Chapter 22 discusses possible reasons for including in the survey a time diary module, which could prompt the respondent to give this complete accounting.

10. Previous LSMS surveys asked detailed questions about in-kind payments, employer characteristics, and working conditions for the respondent's main and secondary jobs during the previous 12 months as well as for his or her main and secondary jobs during the previous 7 days. Requiring respondents to provide details on the 12-month reference period jobs came at some cost—increasing the chance that respondent fatigue would lead to poor data quality in the rest of the questionnaire—and little benefit—as the detail has not been used much by researchers. Looking across rural and urban areas in several LSMS countries, the percentage of individuals with complete 7-day reference period data who reported additional 12-month reference period activities (and whose observations in the dataset were thus put at greater risk of fatigue problems by having to provide detailed 12-month reference period information) was often in the range of 5 to 10 percent, and ran as high as 28 percent.

11. In the Côte d'Ivoire survey the age limit was six years. The percentage of children who reported having done some work during the previous 7 days was approximately 4 percent at the age of seven, 9 percent at the age of eight, 20 percent at the age of ten, 23 percent at the age of twelve, and 40 percent at the age of fourteen.

## References

- Bulow, J., and L. Summers. 1986. “A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination and Keynesian Unemployment.” *Journal of Labor Economics* 4 (3): 376–413.
- Hall, R. E. 1982. “The Importance of Lifetime Jobs in the United States Economy.” *American Economic Review* 72 (4) : 716–24.
- Harris, J., and M. Todaro. 1970. “Migration, Unemployment and Development: A Two-Sector Analysis.” *American Economic Review* 60 (1): 126–142.
- Heckman, J.J., and J. Hotz. 1986. “An Investigation of the Labor Market Earnings of Panamanian Males: Evaluating the Sources of Inequality.” *Journal of Human Resources* 21 (4): 507–42.
- Hussmans, R., F. Mehran, and V. Verma. 1990. *Surveys of Economically Active Population, Employment, Unemployment and Underemployment: An ILO Manual on Concepts and Methods*. Geneva: International Labour Office.
- Lam, D., and R. Schoeni. 1993. “Effects of Family Background on Earnings and Returns to Schooling: Evidence from Brazil.” *Journal of Political Economy* 101 (4): 710–40.

- MaCurdy, T.E. 1985. "Interpreting Empirical Models of Labor Supply in an Intertemporal Framework with Uncertainty." In J.J. Heckman, ed., *Longitudinal Analysis of Labor Market Data*. New York: Cambridge University Press.
- Mroz, T. 1987. "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions." *Econometrica* 55 (4): 765-99.
- Rodgers, W. L., C. Brown, and G. J. Duncan. 1993. "Errors in Survey Reports of Earnings, Hours Worked and Hourly Wages." *Journal of the American Statistical Association* 88 (424): 1208-18.
- Sicherman, N. 1990. "The Measurement of On-the-Job Training." *Journal of Economic and Social Measurement* 16 (4): 221-30.
- Schaffner, J.A. 1999. "Job Stability in Developing and Developed Countries: Evidence from Colombia and the United States." Tufts University, Fletcher School of Law and Diplomacy, Medford, Mass.
- . 1998. "Premiums to Employment in Larger Establishments: Evidence from Peru." *Journal of Development Economics* 55 (1): 81-113.
- . 1997. "The Sensitivity of Wage Equation Estimates for a Developing Country to Changes in Sample Selection Model Specification." Stanford University, Department of Economics, Palo Alto, Cal.
- Stiglitz, J. E. 1974. "Alternative Theories of Wage Determination and Unemployment in LDC's: The Labor Turnover Model." *Quarterly Journal of Economics* 88 (2): 194-227.
- Strauss, J., and D. Thomas. 1995. "Human Resources: Empirical Modeling of Household and Family Decisions." In T. N. Srinivasan and J. R. Behrman, eds., *Handbook of Development Economics*. Vol. 3. Amsterdam: North Holland.

# 10 Anthropometry

*Harold Alderman*

Anthropometry, the measurement of human growth and size, is widely considered to be a non-invasive, inexpensive way to assess the nutritional status of large samples of individuals. By providing information on one dimension of an individual's health status, it can reflect his or her intake of nutrients and morbidity history. These are important dimensions of welfare that can influence the consumption and investment choices of the household of which the individual is a member.

Anthropometry can be used in clinical settings both to make medical diagnoses and to assess whether individuals are eligible to be included in targeted programs.<sup>1</sup> However, the anthropometric measures of nutritional status (AMNS) that are derived from multitopic household surveys differ appreciably from clinical screening techniques, as the clinical techniques are directly linked to the individuals being measured while multitopic household surveys are a research tool based on a representative and anonymous sample. The data derived from these surveys have been used to raise public awareness of particular nutritional issues and to inform the analysis and evaluation of policies aimed at combating the causes of malnutrition (Alderman 1995).

The first section of this chapter reviews the policy issues that can be addressed using anthropometric survey data. The second section deals with what anthropometric data are needed for effective policy analysis. The third section discusses the draft anthropology module (in Volume 3), and the fourth section provides observations and notes about this module.

## **Policy Issues Regarding Anthropometric Survey Data**

Anthropometric measures of nutritional status can usefully augment the limited portrait of living standards that is revealed by the money value of goods and services consumed (UNDP 1990, Ravallion 1993). For example, Steckel (1995) shows how historical patterns in adult height and weight shed light on patterns of economic growth over periods of up to two centuries. In addition, the fact that there is no perfect correlation between AMNS and either national income levels or national income distribution is often used to distinguish countries that are atypical or to motivate research to account for this atypicality. In places such as Sri Lanka or the Indian state of Kerala, the provision of public services has led to higher levels of health than might have been expected given their aggregate level of income or rates of poverty (Anand and Ravallion 1993). On the other hand, nutritional status in some countries has not improved as rapidly as might have been expected given the countries' income growth—perhaps indicating a need to make specific

investments in human resources (Alderman and Garcia 1994).

Moreover, unlike other indicators of living standards for which data are collected only at the household level, anthropometric measures can provide insights into the distribution of resources within the household, on both a gender and a birth order basis, because the data are collected on an individual basis. While the evidence for gender-specific patterns of childhood malnutrition is mixed,<sup>2</sup> patterns and anomalies can both be used to focus attention on questions of intrahousehold allocation.

This section, a review of policy issues, begins with a discussion about using anthropometric data to indicate both the welfare of a population or subpopulation and the success or failure of poverty and health interventions. The section goes on to discuss the use of AMNS to indicate the consequences of malnutrition and analyze the determinants of malnutrition in a population. Finally, the section examines how these data can be used to design more effective interventions to combat malnutrition.

#### *Using Anthropometry to Assess Poverty*

Descriptive statistics on the anthropometry of children are regularly used in poverty assessments and development statistics, usually to show rates of malnutrition. International comparisons are often cited to argue that a given country needs to pay more attention to the health and nutritional status of its population. However, the strength of this argument depends in part on whether anthropometry provides a common way of measuring health and nutritional status that is more accurate than the exchange rates (purchasing parities) used to compare income and poverty data.

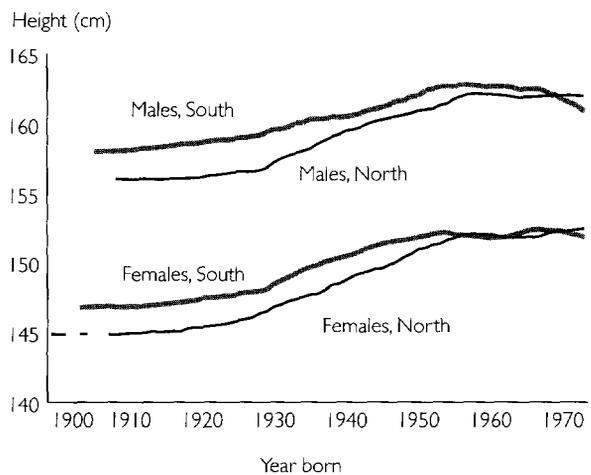
In many studies, aggregate anthropometric measures are presented for children under five years of age. However, more disaggregated analysis of AMNS for this age group can reveal patterns by age—patterns influenced by factors such as birth weight, weaning practices, and exposure to pathogens. This degree of disaggregated reporting can often be found in the results of a survey such as a Demographic and Health Survey, which collects little economic data but more data on child care practices than a household survey would normally collect. Where results are available from a series of repeated surveys, it is possible to compare the status of a cohort of children over time, pos-

sibly indicating the impact of changing economic conditions over short periods.

In cases where the heights of adults are reported, these figures are considered to be an index of economic welfare (Fogel 1994; Steckel 1995). Strauss and Thomas (1998) show that economic trends can be indicated with cross-sectional survey data by plotting the heights of cohort groups. Figure 10.1 presents the usage of this approach for Vietnam. This illustrates both the long-run trend and the leveling off of the trend possibly due to intensification of the civil war. It also indicates a convergence over time between the heights of individuals in the north and the south part of the country. However, because the use of completed heights requires a time lag of 20 years, adult heights tell us little about current economic conditions.

Body mass index (BMI), which is defined as an individual's weight in kilograms divided by the square of the individual's height in meters, is another measure of adult health that can be derived from anthropometric data. BMI is highly correlated with many health-related indicators, including mortality risk (Calle and others 1999; Gibson 1990; Waaler 1984; Fogel 1994). However, unlike most measures of children's nutritional status, BMI values represent increased health risks at both low and high levels. Low BMI can indicate hunger. High BMI indicates obesity and resulting risks of high blood pressure, diabetes, and stroke. Thus, for some countries, the number of adults with high BMI levels may be as indicative of

**Figure 10.1 Height of Adults in Vietnam, 1992–93**



Note: Individuals were categorized as from the North or from the South based on their place of residence at the time of the survey.

Source: Author's calculations from Vietnam Living Standards Survey.

public health problems as the number of adults with low BMI levels.

The same measures of anthropometry that can be used to assess poverty can also be used for targeting interventions or for assessing whether targeting criteria are adequate. For example, it is a straightforward process to use survey data to calculate the percentage of a transfer or the percentage of a subsidized commodity that accrues to households with malnourished members. However, it is less straightforward to determine how many people would have been malnourished had the transfer or subsidy not been in place. Moreover, as Glewwe and van der Gaag (1990) demonstrate, different measures define different individuals or households as poor. Thus, just because a program does not meet its nutritional targets does not necessarily mean that the program is not meeting the objectives that its proponents intended it to meet.

#### *Modeling the Consequences of Malnutrition*

Analysis of AMNS can both demonstrate the consequences of malnutrition and indicate the returns to investments in health. The value of an investment in nutrition is most directly evident if the contribution of malnutrition to other health indicators, such as infant and child mortality, can be determined (Pelletier 1994). However, it is easier to find an association between malnutrition and mortality than to find a causal relationship.

Not only do investments in nutrition directly contribute to human welfare by improving the health of their recipients, they also enhance the efficacy of other government investments. For example, improving children's nutritional status can boost school attendance and enhance students' ability to learn (Pollitt 1990; Behrman 1996). To explore this possibility, it is necessary to know the relationship between nutrition and schooling decisions or outcomes (Glewwe and Jacoby 1995; Alderman and others 1997).

The relationship between nutrition and schooling decisions or outcomes is most easily shown with data from samples specially designed to evaluate the effect of a program.<sup>3</sup> Nonetheless, national, cross-sectional survey data, with suitable controls in the analysis for simultaneity in households' decisionmaking processes, can show the association of nutrition status and schooling decisions or outcomes (Glewwe and Jacoby 1995). These relationships can be explored using either data from a special purpose sample survey designed to

evaluate specific programs or national, cross-sectional survey data and analysis techniques that control for simultaneity in the decisionmaking process.

Similarly, AMNS can be used to show the relationship of nutrition to worker productivity, measured in terms of wages or agricultural output. Studies that use household-level data to show this link include Thomas and Strauss (1997) and Haddad and Bouis (1991). In addition, Fogel (1994) uses historical patterns to back up his claim that improved nutrition accounts for an appreciable share of income growth in the past two centuries.

These studies of the consequences of malnutrition generally require analyses that place nutrition on the right-hand side of regressions. While height is generally taken as predetermined in wage regressions, it is harder to justify treating weight as exogenous in such equations or to claim that any measure of nutrition is not simultaneously determined with schooling choices. This imposes analytical challenges that cross-sectional data only partially address. These challenges are discussed in greater detail below.

#### *Using Anthropometry to Assess Interventions and Policies*

Anthropometric measures of nutritional status can be used to demonstrate the success of a program or policy in reducing malnutrition in a given country. They can show whether there is a need for explicit investments in nutrition to augment the gains that might be expected from a labor-intensive growth strategy or find out whether a proposed transfer program would have an appreciable impact on the nutritional status of the population. Alternatively, they can be used to predict how general changes in economic conditions will affect health status or how expanding access to education can change the population's nutritional status. In all of these cases, it is first necessary to derive from the survey data an estimate of the impact that increased income or improved education have on nutritional status. That parameter is then multiplied by the size of the gain expected in income from growth or from the hypothetical welfare program.<sup>4</sup>

Similarly, causal models of malnutrition can be used to measure how price changes affect malnutrition and, thus, to assess how removing a subsidy or changing the exchange rate will affect it. Analogously, such models can be used to ascertain whether certain types of infrastructure or specific programs contribute to reducing malnutrition (Strauss and Thomas 1995).

Since an individual's nutritional status is the outcome of a complex process of household decisions, data on nutritional status can also be used to make inferences about how a household allocates its resources. For example, Horton (1988) explores how birth order affects malnutrition, providing both specific insights on nutrition and general insights on household resource constraints. Similarly, Thomas (1994) explores the difference in the impact on a child's AMNS of the amount of household income controlled by mothers and by fathers and how this impact differs if the child is a boy or a girl. The study by Thomas presents strong evidence that resources are not fully pooled within a household but rather that some members may retain control over the disposition of resources that they bring into the household.

Another study in which AMNS are used to model general issues of household decisionmaking is Pitt, Rosenzweig, and Hassan (1990). This study uses a weight for height production function to derive an individual's health endowment. The study is longitudinal, meaning that it uses data collected from a number of successive rounds of the same survey. Once the individual health endowment is derived, it is used to indicate how households reduce internal inequality by transferring output from the household members with the highest endowments to members with lower endowments (see also Filmer 1995).

In another study using AMNS, Foster (1995) creates a model for weight changes in Bangladeshi children over several survey rounds. This model is used to indicate how sensitive households are to income shocks caused by flooding. Although the primary contribution of this study is in terms of understanding credit markets, it also provides a useful perspective for understanding child health.

Foster and Rosenzweig (1994) use adult nutrition as an outcome measure in their study of contractual arrangements for workers in the Philippines. Unlike most of the studies cited above, this paper does not aim to improve health policies but rather investigates the incentives to work provided by different labor contracts. However, the underlying hypothesis of the authors' discussion is that work effort is reflected in AMNS. Thus the authors use AMNS as an objective measure of a personal choice that is otherwise difficult to observe.

Many of the studies that use LSMS data to model nutritional outcomes use reduced-form estimations.

These have only exogenous (predetermined) variables on the right-hand side and are less data-intensive than properly modeled production functions. They provide the net impact of the exogenous (predetermined) factors and are often suitable for determining the effect of a particular intervention or service. For example, they can show the net impact of the availability of a prenatal clinic in the community without actually modeling who uses the clinic or the direct impact of people's use of the clinic.

Other studies of the determinants of malnutrition in children adopt production function approaches, which use household decisions about health care to explain health outcomes. These approaches can be used to study the impact of specific behaviors such as smoking, breastfeeding, or using oral rehydration salts. More than most approaches, production function approaches are stable over changing economic circumstances and amenable to making extrapolations beyond existing conditions (Rosenzweig and Schultz 1988). However, these approaches require specific, detailed data on all the elements in the production function. While multitopic household surveys often contain sufficiently detailed estimates of prices to identify the inputs into production functions, the identification process is not always precise. Production function approaches are most informative when they have well-defined inputs. Many key inputs into nutrition are not collected in the short LSMS questionnaire, including individual nutrient intakes, breastfeeding practices, the timing and frequency of feeding, and the use of oral rehydration salts.

### **Data Needs**

This section discusses what data are needed for anthropometric analysis and the extent to which data from LSMS-type surveys can fill these analytical needs.

#### *What Measures for What Ages?*

Since the publication of Waterlow and others (1977), a distinction has generally been made between stunting (low height for age), which is considered a measure of long-term or chronic malnutrition, and wasting (low weight for height), which is considered a measure of acute malnutrition. A third category, weight for age, contains some of the information from both these measures and is still commonly reported.<sup>5</sup> There is no apparent correlation between levels of stunting and

levels of wasting in a population (Victora 1992). This is a bit of a puzzle since a cumulative measure of health such as height should reflect the sum over time of short-term health status, which wasting should measure. There have been few attempts to use household survey data to solve this puzzle.

**HEIGHT FOR AGE.** Stunting (low height for age) is an indicator of long-term or chronic malnutrition. All LSMS surveys to date that have collected AMNS have included the height for age measure, and all analyses of AMNS data have discussed height for age results. Therefore, it is not necessary to address here whether height for age should be included in modules on anthropometrics.

The measurement of height will be further discussed below in the context of age-specific concerns. However, it should be noted that it is strongly recommended not to seek self-reported heights or to ask parents to indicate the heights (or weights) of their children. Strauss and Thomas (1996) report that even with high-quality data from the United States there is a systematic bias in heights reported by parents. Low-income households underestimate their children's heights more than do higher-income parents—which can result in an exaggeration of the impact of income on height. Strauss and Thomas also indicate that height variance appears much greater when it is self-reported.

**WEIGHT FOR HEIGHT.** Wasting (low weight for height) is, in principle, an indicator of acute malnutrition and thus may be a sensitive indicator of short-term response to changing conditions. However, this is not necessarily desirable in cross-sectional analysis. Multipurpose household surveys are not generally designed to differentiate short-term fluctuations from long-term conditions. The variation in prices and infrastructure in a cross-sectional data set has short-term and long-term components that are hard to distinguish from each other. To overcome this and make the most of the information on short-term nutritional status contained in weight for height measures, it may be necessary to adopt panel approaches.

Weight for height measures probably work best when there are repeated measurements over relatively short periods of time. One of the few studies to explicitly choose weight for height for its sensitivity to short-term health (Pitt, Rosenzweig, and Hassan 1990) uses repeated observations of this measure as

well as of two other measures of nutritional status (skin folds and upper arm circumference). Foster (1995) also specifically examines short-term nutritional status, but he uses changes in weight, not weight for height.

The importance of data from the community modules is highlighted by comparing the results in Tables 10.1 and 10.2. These tables show that regressions using height for age or weight for height as a dependent variable fail to explain much of the variance in the sample, as indicated by the low values of  $R^2$  unless cluster fixed effects—which capture the effects of community infrastructure—are included. (Cluster fixed effects will be discussed later in this chapter.) Moreover, the coefficients on income are quite different between the weight for height and height for age regressions. In two cases the signs on the variables for the impact of maternal education differ depending on which measure of nutritional status is used. While differences in results do not, by themselves, indicate that one measure is more accurate than the other, the regressors used to produce Table 10.2 were not tailored to the measurement of acute conditions, so it is presumed that the height for age long-run measure is more suitable for the analysis reported.

**WEIGHT FOR AGE.** Although weight for age is still a commonly used indicator, its use has declined, in part because it is viewed as unable to distinguish between chronic and acute malnutrition. It may, however, regain some favor with wider use of accurate battery or solar-powered digital scales. Such scales are particularly useful for young children, whose weights can be calculated conveniently by first weighing the mother or caretaker with her child and then weighing her alone. Often the scale performs this calculation, and hence the subtraction does not introduce additional error. Nevertheless, relatively few analyses of weight for age have been done linking this measure with economic data from LSMS surveys. Table 10.3, which uses the same sample and models as Tables 10.1 and 10.2, indicates that the results from analyses of weight for age data correspond more closely to results based on height for age than to results based on weight for height. Even if weight for age provides little information beyond that already provided by height for age, there is little additional cost to collecting weights along with heights. Moreover, as discussed below, in a number of analyses, weights can be used as additional instrumental variables to increase the accuracy of height coefficients.

**Table 10.1 Sensitivity of Coefficients in Three Models to Explain Height for Age, Selected LSMS Surveys**

	Coefficient logarithm of income	Coefficient of maternal education	Coefficient of paternal education	F test for price coefficients	R <sup>2</sup>
<i>Vietnam (1992–93)</i>					
Community fixed effects model	0.232 (3.132)	0.018 (1.722)	0.016 (1.726)		0.283
Model without fixed effects	0.373 (6.422)	-0.003 (0.273)	0.008 (0.870)		0.225
Model without fixed effects but including commodity prices	0.284 (4.614)	0.005 (0.475)	0.013 (1.496)	3.59 [6, 2,594]	0.232
<i>South Africa (1994)</i>					
Community fixed effects model	0.273 (5.517)	0.012 (1.282)	0.007 (0.826)		0.239
Model without fixed effects	0.355 (8.276)	0.015 (1.565)	0.002 (0.185)		0.125
Model without fixed effects but including commodity prices	0.337 (7.661)	0.015 (1.582)	0.002 (0.207)	2.51 [6, 3,277]	0.129
<i>Pakistan (1991)</i>					
Community fixed effects model	0.189 (2.641)	0.057 (5.247)	0.012 (1.638)		0.304
Model without fixed effects	0.325 (5.390)	0.060 (6.328)	0.021 (3.003)		0.142
Model without fixed effects but including commodity prices	0.382 (5.968)	0.061 (6.253)	0.01 (2.483)	5.24 [7, 3,749]	0.148
<i>Morocco (1991)</i>					
Community fixed effects model	0.239 (2.168)	0.038 (2.137)	0.005 (0.150)		0.338
Model without fixed effects	0.296 (3.172)	0.040 (2.423)	0.006 (0.153)		0.192

Note: Numbers in parentheses are t-statistics and numbers in brackets are degrees of freedom. Regressions include variables for age, gender, and the interactions between age and gender. Parental heights are modeled where available. Regressions for Vietnam additionally include variables for race.

Source: Author's calculations from LSMS data for each country.

**ARM CIRCUMFERENCE.** Mid-upper arm circumference is a measure that gauges both fat reserves and muscle mass. In contrast to the bulky scales and measuring boards needed to collect other measurements, it requires little equipment other than a calibrated tape measure (Zerfas 1991). Thus, in places where interviewers have to carry anthropometric equipment long distances to sample households, it may be more practical to measure mid-upper arm circumference than to measure weight and height. Moreover, in longitudinal surveys mid-upper arm circumference has at least as strong a correlation with subsequent mortality as do other measures (Vella and others 1993).

The World Health Organization (WHO; 1995) recommends against using unadjusted mid-upper arm circumference because it is not age-independent. However, conceptually it is a simple matter to adjust the measurement for age using the data provided in WHO (1995). Since this standardization is routinely done for heights and weights, by itself it does not seem

a compelling reason to prefer one measure over another. Indeed, Zerfas (1991) claims that mid-upper arm circumference adjusted for age is relatively insensitive to errors in recording age. In practice, standardizing mid-upper arm circumference for age is less convenient than standardizing height or weight for age, since no prepackaged software that makes the adjustment easily is yet available.

A few studies have indicated that mid-upper arm circumference is not strongly correlated with weight for height (WHO 1995). However, the absence of such a correlation does not seem to rule out the validity of mid-upper arm circumference as a measure of nutritional status; height for age, also not strongly correlated with weight for height, is widely used as such a measure. So far Vietnam is the only country where a LSMS survey has included mid-upper arm circumference. In Table 10.4, both individual correlations of anthropometric measure for children in Vietnam and the correlation of rates in the 148 sample clusters are

**Table 10.2 Sensitivity of Coefficients in Three Models to Explain Weight for Height, Selected LSMS Surveys**

	Coefficient logarithm of income	Coefficient of maternal education	Coefficient of paternal education	F test for price coefficients	R <sup>2</sup>
<i>Vietnam (1992–93)</i>					
Community fixed effects model	0.290 (0.701)	–0.073 (1.363)	–0.074 (1.800)		0.128
Model without fixed effects	0.208 (0.541)	–0.037 (1.086)	–0.058 (1.280)		0.080
Model without fixed effects but including commodity prices	0.383 (0.953)	–0.095 (1.106)	–0.066 (1.506)	0.96 [6, 2,594]	0.080
<i>South Africa (1994)</i>					
Community fixed effects model	0.085 (1.627)	0.001 (0.093)	0.004 (0.570)		0.250
Model without fixed effects	0.047 (1.062)	0.005 (0.602)	0.004 (0.415)		0.053
Model without fixed effects but including commodity prices	0.023 (0.512)	0.01 (1.088)	0.008 (0.889)	13.28 [6, 3,277]	0.077
<i>Pakistan (1991)</i>					
Community fixed effects model	0.155 (2.547)	0.001 (0.156)	0.009 (1.454)		0.256
Model without fixed effects	0.198 (3.801)	0.001 (1.453)	0.009 (0.505)		0.025
Model without fixed effects but including commodity prices	0.165 (2.981)	–0.006 (0.722)	0.003 (0.532)	8.23 7, 3,749]	0.044
<i>Morocco (1991)</i>					
Community fixed effects model	0.072 (0.836)	–0.041 (3.055)	–0.008 (0.301)		0.267
Model without fixed effects	0.219 (2.988)	–0.027 (2.158)	0.004 (0.141)		0.018

Note: Numbers in parentheses are t-statistics and numbers in brackets are degrees of freedom. Regressions include variables for age, gender, and the interactions between age and gender. Regressions for Vietnam also include variables for race.

Source: Author's calculations from LSMS data for each country.

reported. While cluster rates are based on very small numbers and thus only roughly indicate similar correspondences in larger samples, available statistics show that rates of low weight for age are poorly correlated with other measures at the cluster level. Thus the table reinforces the view that weight for height reflects a very different dimension of community nutrition than is reflected in height for age or weight for age. Mid-upper arm circumference may be a more short-term measure than height, as muscle mass and fat reserves can both decrease as well as increase. Height, however, is as much a measure of past nutritional shocks as of current circumstances, and does not appear to be as closely related to weight for height as it is to other measures.

There is little evidence on how mid-upper arm circumference correlates with the economic variables collected in integrated household surveys, since this measure has rarely been collected in LSMS surveys. Nor is mid-upper arm circumference often used in studies linking economic data with nutrition. Pitt,

Rosenzweig, and Hassan (1990) use this measure, not to study such links but in a statistical technique that requires an over-identifying instrument. Table 10.5 repeats three of the regressions reported for Vietnam in Tables 10.1–10.3, using mid-upper arm circumference standardized for age as the dependent variable. The coefficients of income and maternal education in the preferred fixed effects estimates are moderately similar to the corresponding regression with height for age as a dependent variable. The same bias can be seen in the coefficient of income relative to the fixed effect estimate observed in other AMNS.

It is not possible to take these comparisons further since the other LSMS data sets do not include mid-upper arm circumference as a variable (although a survey currently being prepared in Paraguay will do so). However, many of the analyses that have been carried out to date with other LSMS data could easily have been undertaken using mid-upper arm circumference data as well. Given this possibility, the strongest argument against the widespread use of mid-upper arm

**Table 10.3 Sensitivity of Coefficients in Three Models to Explain Weight for Age, Selected LSMS Surveys**

	Coefficient logarithm of income	Coefficient of maternal education	Coefficient of paternal education	F test for price coefficients	R <sup>2</sup>
<i>Vietnam (1992–93)</i>					
Community fixed effects model	0.167 (0.370)	0.019 (2.517)	0.001 (0.011)		0.412
Model without fixed effects	0.198 (4.036)	-0.01 (1.488)	-0.002 (0.247)		0.351
Model without fixed effects but including commodity prices	0.154 (3.592)	0.1017 (2.403)	0.002 (0.253)	4.36 [6, 2,594]	0.357
<i>South Africa (1994)</i>					
Community fixed effects model	0.223 (4.740)	0.008 (0.944)	0.010 (1.267)		0.277
Model without fixed effects	0.239 (6.046)	0.013 (1.565)	0.004 (0.587)		0.128
Model without fixed effects but including commodity prices	0.210 (5.135)	0.016 (2.011)	0.008 (1.028)	8.89 [6, 3,277]	0.144
<i>Pakistan (1991)</i>					
Community fixed effects model	0.220 (3.958)	0.033 (3.952)	0.016 (2.955)		0.347
Model without fixed effects	0.330 (7.170)	0.025 (3.134)	0.016 (0.016)		0.146
Model without fixed effects but including commodity prices	0.335 (6.718)	0.030 (3.677)	0.014 (2.558)	7.44 [7, 3,749]	0.158
<i>Morocco (1991)</i>					
Community fixed effects model	0.215 (2.463)	-0.014 (1.118)	-0.009 (0.358)		0.346
Model without fixed effects	0.366 (5.163)	-0.002 (0.191)	0.002 (0.088)		0.180

Note: Numbers in parentheses are t-statistics and number in brackets are degrees of freedom. Regressions include variables for age, gender, and the interactions between age and gender. Regressions for Vietnam also include variables for race.

Source: Estimate from LSMS data.

**Table 10.4 Correlation of Anthropometric Measures for Children 0–60 Months, Vietnam, 1992–93**

Variable	Mean value	Rate of malnutrition	Correlation over individuals				Correlation of rates over clusters			
			Height for age	Weight for age	Weight for height	Mid-upper arm circumference	Height for age	Weight for age	Weight for height	Mid-upper arm circumference
Height for age	-1.89	49.9	1.00	—	—	—	1.00	—	—	—
Weight for age	-1.62	40.8	0.75	1.00	—	—	0.68	1.00	—	—
Weight for height	-0.59	5.8	-0.02	0.62	1.00	—	0.07	0.21	1.00	—
Mid-upper arm circumference	-1.84	49.7	0.42	0.67	0.50	1.00	0.42	0.53	0.21	1.00

— These correlations are the same as the corresponding correlations below the diagonal of this 4x4 matrix.

Source: Estimate from LSMS data.

**Table 10.5 Three Models to Explain Mid-Upper Arm Circumference, Vietnam, 1992–93**

	Coefficient logarithm of income	Coefficient of maternal education	Coefficient of paternal education	F test for price coefficients	R <sup>2</sup>
Community fixed effects model	0.183 (3.144)	0.026 (2.841)	0.007 (0.922)		0.384
Model without fixed effects	0.351 (7.394)	0.006 (0.689)	-0.008 (0.397)		0.277
Model without fixed effects but including commodity prices	0.251 (4.931)	0.012 (1.473)	0.000 (0.006)	4.42 [6, 2,594]	0.284

Note: Numbers in parentheses are t-statistics and numbers in brackets are degrees of freedom. Regressions include variables for age, gender, and interactions of age and gender.

Source: Estimate from LSMS data.

circumference data appears to be that it has not been used widely so far. In other words, this measure cannot be used as a poverty indicator because there are too few examples for cross-country comparisons to be made. And mid-upper arm circumference statistics are rarely used to inform policy probably only because the indicator is relatively new. Thus it would be appropriate in the future to more frequently include this indicator in household survey questionnaires .

**BIRTH WEIGHTS.** While it is well known that birth weights are a strong predictor of the subsequent size of children (Adair 1989), it is generally not possible for multitopic household surveys to gather birth weight information in any systematic way. Birth weight data are occasionally recorded on health cards that are issued by clinics and kept by a baby's parents, but in most developing countries the percentage of children born in a clinic is usually too small to provide a reliable indicator without correcting for sample selection. In countries where these data are available for a significant portion of the population, it would be advisable to take advantage of this opportunity to collect them. However, in most cases, designers of LSMS surveys should assume that this will not be the case.

**MEASURING YOUNG CHILDREN.** AMNS collection is most important for preschool children since these children are particularly at risk of malnutrition. However, certain difficulties arise when measuring the nutritional status of very young children. Because children under 2 years of age are too young to stand properly, their length must be measured while they are lying down. The international reference tables accommodate this difference in measurement techniques (Dibley and others 1987).

There are also some practical considerations regarding the measurement of very young children. Cultural taboos related to the evil eye and negative feelings regarding strangers visiting a child are most common in the case of the very young. In five data sets investigated—Ghana, Morocco, Pakistan, South Africa, and Vietnam—the probability of a missing height measurement for a preschool child declines significantly with age, with most of this decline coming between the first and second year of life. In some cultures, dead children are buried on boards that resemble the measuring boards used by interviewers, which makes parents understandably reluctant to see their

live children lying on these measuring boards. Other considerations in measuring very young children are the timing of the most recent feeding or the fact that parents may not allow their child's swaddling or diapers to be removed. While the failure to remove clothing can affect the recorded weights of older children too, clothing represents a higher percentage of the weight of a very young child. Also, while misreporting of age can throw off the interpretation of the anthropometric measurements of all children, it is especially problematic for very young children; because very young children grow extremely rapidly, even a small error in assessing their age can result in a significant error in assessing their nutritional status.

These problems justify treating children younger than one year old differently than slightly older children, in analysis as well as in data collection. However, the problems do not justify excluding younger children from the sample for which anthropometric data are collected. Indeed, given the long-lasting impact of early childhood nutrition, identifying and understanding the determinants of nutrition for this age group is one of the most important tasks for analysts.

At times a lack of resources or technical capacity makes it difficult to collect weight and height information. In such cases it is sometimes suggested that the problem be addressed by measuring only children under the age of 36 months. The proponents of this solution justify it on the grounds that peak vulnerability occurs in these years. However, while this may be an appropriate way to prioritize in clinical work, it is not as appropriate in survey work because the costs of locating other, slightly older, children are not high. The key issue for determining the appropriate age cutoff is not how easy it is to take the relevant measurements but international comparability and correlation with observable prices and infrastructure.

**MEASURING ADOLESCENTS.** There is little experience with collecting data on individuals between the ages of 6 and 20. At the low end of this range, these data differ only slightly from data on younger children and can be useful, for example, to explain why some children are enrolled in school at the age of six and others are not. However, by the time a child reaches the prepuberty growth spurt (at approximately age 9), analyzing his or her nutritional status becomes more problematic. While international growth reference curves are available for the height of children up to age

18, they are only available for weight for height of girls up to age 10 and boys up to age 11. Moreover, little work has been done on assessing international patterns or ethnic differences within a given country in the timing of the onset of puberty.

Thus the value of collecting anthropometry for adolescents in household surveys is unproven. However, two arguments can be made in favor of including it in multitopic household surveys. First, increasing the amount of data available on this age group may indicate the data's usefulness with high relative returns, particularly in the case of panel studies that might shed light on how nutrition and schooling affect each other or on whether adolescent children can overcome stunting that occurred when they were younger (Golden 1994; Martorell and others 1992). Second, if the cost of collecting anthropometric data largely consists of the fixed cost of finding the household, including adolescents when measurements are being taken of preschool children and adults is relatively inexpensive.

**MEASURING ADULTS.** Since rates of childhood malnutrition are a more widely monitored indicator of welfare than adult BMI or height, measuring adults is often regarded as of secondary interest for policymaking purposes. Yet measuring adults is useful both for analyzing policy (for example, assessing the impact of nutrition on labor productivity) and for understanding the determinants of the nutritional status of children and other members of the family. Cross-tabulations or multivariate regressions that explain child nutrition may be misleading if adult heights are not available, since parental, especially maternal, heights are significant explanatory variables in many studies of children's nutrition. Moreover, adult stature may correlate with the income and welfare of populations. And adult heights and body mass indexes can be used to provide information on the long-term returns to investments in nutrition.

In multitopic household surveys, stature and weight of adults are less frequently collected than stature and weight of children. One possible reason for this is that it may be more difficult to collect such measures from adults. If children are weighed using hanging scales, the measurement of adults requires that the team carry an additional, somewhat less portable, scale. However, as mentioned above, digital scales are currently available that are accurate for both children and adults.

Another problem with measuring adults is that it may be more difficult to find all adults in the vicinity of the household during the time of the interview. The better educated (as well as healthier) adults are liable to hold jobs that prevent them from being present when the field team visits the respondents' home. Or these adults' jobs may have caused them to migrate. There is a gender bias in availability of adults to be measured; while only 4.7 percent of mothers were missing in the Morocco sample, over 30 percent of fathers were. Even in the Vietnam LSMS—a data set with a remarkably low number of missing measurements overall—more than five times as many fathers as mothers were not measured. Missing height measurements were especially high for males between 20 and 25; in this age group over 10 percent of the sample was missing.

As a result of the relative scarcity of adult measurements in the field, there is less existing evidence on adult health, as represented by BMI, than there is on child malnutrition. (Studies of adult BMI are included in Thomas, Lavy, and Strauss 1996 and Lavy, Thomas, and de Vreyer 1996; a few studies, such as Pitt, Rosenzweig, and Hassan 1990, pool children and adults.) Adult BMI models are conceptually similar to models of child nutrition, although children may be more passive recipients of care than adults. However, some additional factors can be measured for adults. Total family income (or expenditure) is more clearly simultaneously determined with adult BMI than it is with nutrition of children. In a related vein, Higgins and Alderman (1997) indicate that since the energy outlay of labor affects BMI, the failure to include a measure of time allocation or energy intensity in a study of adult health may result in biased income coefficients. And pregnancy and lactation need to be included in models of female BMI.

Table 10.6 shows the differences in income and education coefficients derived from three LSMS surveys with and without parents' heights, using the cluster fixed effect models (with relevant coefficients repeated for convenience).<sup>6</sup> It is apparent that in both Vietnam and Morocco excluding the parental height variable leads to an appreciable bias in the coefficient of per capita expenditures, while the education coefficients do not seem to be affected. In the case of Pakistan, for which only mothers' heights are available, the difference in the coefficient of income is relatively small when the mother's height is excluded.

**Table 10.6 Effect of Omitting Parental Height in Three Models of Height for Age, Selected LSMS Surveys**

	Coefficient logarithm of income	Coefficient of maternal education	Coefficient of paternal education	R <sup>2</sup>
<i>Vietnam (1992–93)</i>				
Community fixed effects model with both parents' heights	0.232 (3.132)	0.018 (1.722)	0.016 (1.726)	0.283
Community fixed effects model with mother's height only	0.257 (3.495)	0.018 (1.707)	0.017 (1.792)	0.276
Community fixed effects model excluding parents' heights	0.283 (3.816)	0.019 (1.791)	0.017 (1.739)	0.259
<i>Pakistan (1991)</i>				
Community fixed effects model with mother's height only	0.189 (2.641)	0.057 (5.247)	0.012 (1.638)	0.304
Community fixed effects model excluding parents' heights	0.204 (2.834)	0.057 (2.196)	0.015 (2.098)	0.299
<i>Morocco (1991)</i>				
Community fixed effects model with both parents' heights	0.239 (2.168)	0.038 (2.137)	0.005 (0.150)	0.338
Community fixed effects model with mother's height only	0.253 (2.277)	0.036 (2.025)	0.004 (0.112)	0.326
Community fixed effects model excluding parents' heights	0.358 (3.263)	0.039 (2.136)	0.004 (0.116)	0.305

Note: Numbers in parentheses are t-statistics. Regressions include variables for age, gender, and the interactions between age and gender. Regressions for Vietnam also include variables for race.

Source: Estimated from LSMS data.

The inclusion of parents' heights does not appear to be as important in preventing bias in the mid-upper arm circumference regression. When heights were omitted in the regression reported in the first row of Table 10.5, the coefficient of income rose only to 0.191. Using parental arm circumference as a regressor is not appropriate since, unlike height, parental arm circumference is affected by the same factors that determine child arm circumference and thus will contribute to biased results.<sup>7</sup>

#### *What Data Complement Anthropometric Data?*

Anthropometric data alone are not sufficient to analyze a number of key nutritional issues. Therefore, the anthropometric data collected in most multitopic household surveys need to be supplemented by data either from other parts of the survey or from outside sources. The LSMS surveys have an advantage over many other sources of anthropometric data in that they gather information on other aspects of the household's well-being, including total expenditure and educational attainment.<sup>8</sup> As is increasingly recognized by epidemiologists and nutritionists as well as economists, this feature makes it possible to be discriminating in making causal statements about observed correlations (Briscoe, Akin, and Guilkey 1990). This subsection of the paper reviews what kinds of data are needed to complement the AMNS.

**OTHER HEALTH INDICATORS.** Many aspects of micronutrient status or deficiencies that affect mortality and productivity are not revealed by anthropometric measurements. Therefore, to do a full study of the nutritional status of a population, it is necessary to collect data on the levels of vitamin A or iron in the blood. Collecting this data is not usually cumbersome or costly, as there are some cheap and convenient field collection methods. Iron status, for example, can be indicated using a blood sample taken by means of a single pinprick. However, experience suggests that research protocols are not always strictly observed by all survey interviewers in the field. The consequences of such deviations can be particularly serious when blood is being collected. Moreover, informed consent becomes both more important and more difficult to obtain when even marginally invasive techniques are used. Thus the collection of blood samples is more suited to specialized clinical studies than to living standards surveys such as the LSMS.

The consumption of micronutrients is, in general, hard to indicate in multitopic household surveys, in part because most food-related data are collected at the household level and thus have to be aggregated over all household members and all meals. This is important as the preparation of the food and the timing of its consumption influence the absorption of

micronutrients. However, it is possible to test for the iodization of salt by asking households to provide a small amount of salt to be tested in the field (as was done in Peru in 1995) so that researchers can ascertain within reasonable bounds whether the household is receiving an adequate amount of iodine.

Some inferences on nutrient intake can be made using food recall data. To provide this data, respondents attempt to remember what food the household has eaten during, say, the past 24 hours. Although LSMS surveys have not traditionally included food recall questions, it would be feasible to include them in a multitopic household survey. Moreover, these data need not be collected on a household basis; they can also be collected on a per meal basis or for an individual or subset of individuals in the household. Indeed, food recall data collected on an individual basis may measure household consumption more accurately than do food expenditures (Bouis and Haddad 1992). If so, they may help identify the contributions of household resources and public health measures toward preventing malnutrition.

A number of new techniques for assessing adult health have recently been introduced into household surveys. These include self-reported assessments of respondents' ability to perform daily activities ranging from walking one kilometer or lifting heavy objects to simple activities such as eating, bathing, or dressing. Other techniques include measuring respondents' lung capacity. The various techniques are discussed at length in Chapter 8.

It is possible to include the anthropometric measures discussed in this chapter in the health module rather than in an independent module. Conceptually, it matters little which choice is made. The main issues that should determine this have to do with the organization of the fieldwork. LSMS surveys often involve anthropometrists who collect only anthropometric data, are well trained in the technique, and are each issued a set of field equipment. If the anthropometric module were incorporated into the health module, either anthropometrists would need to be trained in general interview techniques or interviewers would need to be trained (and equipped) for anthropometry. Linking the two modules might make it easier to use the time of all the members of the field team more efficiently, reducing the chance that one team member would have to wait while his or her colleagues completed the other aspects of the survey. Moreover, put-

ting objective measurements in the same module as questions that require reported answers makes it clearer which data have been reported by the individual him or herself and which have been reported by other family members.

The fertility modules within LSMS surveys have often included questions on breastfeeding and immunization that are important for anthropometric analyses (see Chapter 15). The general practice is to put these questions in the fertility module because they refer not only to living children but also to any children who have died. If a survey does not include a fertility module, it would be useful to include questions on breastfeeding and immunization within the anthropometry module.

A simple yes or no question is seldom sufficiently informative about breastfeeding; interviewers must ask when the breastfeeding began and ended, and whether the child was exclusively breastfed during this period. As the phrase "exclusive breastfeeding" is ambiguous, the wording of the question should distinguish among various means of providing fluids (both milk and water) and elicit the date on which the child was first introduced to solid foods. For times when a child was not breastfed, a distinction should be made between cup and spoon feeding and bottle feeding. Often questions about breastfeeding also ask whether colostrum was given to the child.

**INCOME AND CONSUMPTION.** Consumption questions are included in all multitopic household surveys, while income questions are included in most such surveys. When consumption and income data are analyzed in combination with anthropometric measures of nutritional status, a few issues need to be considered.

Some difficulties may arise if researchers are interested in calculating the availability of nutrients (and if food consumption recall is not included in the module). Generally, food consumption in LSMS surveys is recorded in terms of expenditures. Using community prices to calculate quantities of goods consumed by a household requires researchers to make a tacit assumption that no differences in quality exist. To the degree that prices vary by income level—as they might if a household can choose a commodity from a range of similar commodities of different qualities (for example, with grains or meat)—using average prices will underestimate the quantities obtained by low-income consumers. Recent analysis has shown that the quality

range is limited in the case of commodities that are defined more precisely—rice, for example, rather than the broader category of grains. The bias in measurement of calories will thus be smaller if the consumption module lists individual foods rather than classes of foods (Subramanian and Deaton 1996; Alderman in Lipton and van der Gaag 1993). Nevertheless, if there is a possibility that prices differ by household characteristics—for example, with poor households buying in small quantities and thus unable to get bulk discounts—the module on food consumption should ask about the most recent price paid by the household.

Some LSMS surveys have inquired about households' recent purchases of food, the number of months households rely on home production, and households' average consumption from home production. Such inquiries do not allow an accurate assessment of current consumption because they do not indicate if home production was being consumed during the period covered in the survey recall. However, most LSMS surveys have collected data on purchases for and consumption from production (as well as in-kind wages), using the same recall period in both cases.

Income and expenditures are often measured imperfectly, in some cases leading to underestimation of the effect of income or expenditures in regressions of the determinants of nutritional status. Chapters 5, 11, and 17 in this book discuss ways to improve the measurement of income and consumption. Yet even without improving income and consumption measurements, it is possible to correct for errors in income and consumption using other information in the multipotopic survey.

The main way to do this is to use the data on household assets from various parts of the questionnaire (such as the modules on durable goods, housing, household enterprises, agriculture, and savings) to apply instrumental variables methodologies. These techniques can also help overcome another problem—the biases in the regressions of the determinants of nutritional status sometimes introduced by income or income proxies due to simultaneous choices. Not only is adult health likely to be dependent on income, income may also reflect the health (and size) of the household members. While such reverse causality is less obvious in the case of the nutrition of young children, it can be hypothesized that household earnings decrease when a child is ill because adult members of

the household must stop their income-earning activities to care for the sick child.

**COMMUNITY DATA.** Community data are very important for nutritional analysis. Yet they remain a part of data collection that needs to be strengthened. (See Chapter 13 for a broader discussion of collecting community data.)

In the community questionnaire, LSMS surveys have typically collected a range of data pertaining to nutritional status, including prices of food and medicine and distance to clinics and doctors. Alternatively, summary variables have been constructed for these indicators by taking the means or medians of the observations (from households within a sample cluster) from various parts of the household questionnaire. However, more specialized questions are needed to monitor disease pathogens or prevalence rates or to gauge the quality of services at health centers. Also, without measuring the availability or rationing of services, the data on prices of medicine may be misleading.

The importance of community data can be seen in the fact that their availability changes the estimated impact of other factors such as income and education. If wealth correlates with the availability of services that are excluded from the regression, some of the effects attributed to income or education may in fact be due to these services. Table 10.1 uses four LSMS data sets to give some indication of this bias. The regressions in Table 10.1 are run with different specifications. One model includes community fixed effects—community means of both observed and unobserved variables. A second model, the conventional model, does not include community information. The coefficients of per capita income in models that regress children's standardized heights for age on a number of regressors in the two models are dramatically different. If the fixed effects models are considered to be unbiased, the bias in models without fixed effects can be as high as 70 percent (as in the case of Pakistan). In other cases the bias is smaller, but these models include race and rural-urban distinctions that may also proxy for service availability. (These results differ somewhat from those reported in Gertler, Glewwe, and Ponce 1998, as the age grouping differs.) In the examples in Table 10.1 the bias on income is always upwards, as it should be if the availability of types of social infrastructure that improve nutrition are positively correlated with income. Thus it is important to include questions on infrastructure in the survey.

The relative magnitudes of the  $R^2$  values of the fixed effects regressions in Table 10.1 also indicate that a great deal of information is shared among the clusters. The cluster fixed effects approach, however, does not make it possible to infer much beyond the presence of some common factors such as prices, altitude, disease vectors, health and sanitation infrastructure, the quality of staff and management in the local clinics, and information and preferences shared among households. More likely, the cluster effects reflect a combination of these common elements.

**COMMUNITY DATA ON HEALTH SERVICES.** LSMS surveys have not usually contained much information on health infrastructure, which has made it difficult for analysts to use LSMS data to devise specific policy recommendations in this area (see Chapter 8). Although some attempts have been made to measure access (usually as defined by distance) to health services, so far only a few multitopic surveys have sought information on the quality of health services. The main reason for this omission is that quality is often difficult to define and clear and accurate information on quality may be hard to get from individual respondents (unlike, say, information on prices or distances). Moreover, for modeling purposes, information is needed not only on the facilities used by the household but also on facilities the household could have chosen to use. Such information is particularly problematic in urban areas, where there may be a range of facilities from which households can choose.

Two studies of nutritional status—Thomas, Lavy, and Strauss (1996) and Lavy, Thomas, and de Vreyer (1996)—study the determinants of malnutrition using data from special modules that contain measures of service quality. Even with such modules, however, quality (or even availability) measures must be handled judiciously. For example, Thomas, Lavy, and Strauss (1996) compare two measures of staffing in Côte d'Ivoire: the number of staff members listed in official records and the number present in the 24 hours preceding the interview. While the actual number of doctors present was found to affect child health favorably and significantly, the number on the books appeared to have no impact on child health. Failing to take this into account may mean that incorrect inferences will be made about which policies are most effective.

In addition (or as an alternative) to collecting better data on the availability and quality of infrastructure,

it may occasionally be possible to link multitopic survey data with other databases such as geographic information systems. For example, efforts are currently underway to link the South African LSMS to the South African Health Management Database System (GIS), a comprehensive database of the country's health facilities and these facilities' staff. In general, however, GIS databases contain limited information on staffing and even less on the quality of services.

**COMMUNITY DATA ON PRICES.** Various policy measures that influence prices (of food, of drugs, or of medical services) also indirectly influence nutrition because the levels of these prices determine what products and services people can afford. A number of studies have used LSMS data to determine how consumers respond to changes in the price of services. These studies have measured how price changes affect demand for health care and the use of alternative service providers, rather than measuring impacts on health status. Nevertheless, anthropometric data from multitopic surveys can also be used to estimate the net impact of prices or service availability on nutrition.

While the prices and availability of services often have an effect on a population's nutritional status, including these factors in the regression does not necessarily mean that there are no biases in the coefficient of income. For example, the coefficients of income in the third row for each country in Table 10.1 differ appreciably from the community fixed effects estimates. This may be because the information on prices and clinic availability that has been collected in most LSMS surveys to date explains only a small share of the cluster-specific information. (An alternative way of looking at this issue is to regress the residuals from regressions similar to those in the third row for each country in Table 10.1 against the cluster fixed effects. When this is done, the F tests for the regressions are significant at conventional levels of significance.)

Moreover, the estimated impact of service availability or prices can be biased if no account is taken of differences in service quality. The measures of travel time contained in short LSMS data sets often do not indicate the quality of the services available at the destination (Strauss 1990). If, for example, higher prices or greater distances are associated with better-quality services, it is likely that the estimated impact of prices on the utilization of health care services or on nutritional status will be understated.

**COMMUNITY DATA ON SANITATION.** As in the case of health infrastructure, most LSMS surveys have not collected enough information on sanitation infrastructure to allow researchers to carry out nutritional analyses. Although it can be hypothesized that the quality and quantity of water delivery and waste removal affect nutrition, most household survey data allow for only broad testing of such hypotheses. This is sometimes done from housing modules by deriving dummy variables that indicate the source of water or the type of latrine; these analyses generally highlight the importance of sanitation. While such dummy variables are not included in the example presented in Table 10.1, including them would not appreciably close the gap in the coefficients of expenditures or in the  $R^2$  values. As with health care, the quality of such services is often hard to gauge unless the survey includes specific modules covering service providers or the analysis is carried out using additional data sources. Chapter 14 contains a discussion of how to enhance the traditional treatment of water and sanitation in LSMS surveys.

**PANEL DATA ON NUTRITION.** There are several types of analysis that would become possible with or could be enhanced by panel data on nutrition. If nutritional status is considered a stock variable and income is considered a flow variable, a regression of height on income really measures the cumulative impact on the stock of height of the flow of income over time. A model that examines flow (growth in height) regressed on other flows, such as income, might be more informative, but growth in height could only be calculated by taking the difference in height between two points in time in a panel data set. One of the few economic studies that model growth (in this case, growth in weight) is Foster (1995), a study of credit markets.

Panel data can be particularly useful for modeling a child's response to interventions. This response is often age-specific (Lutter and others 1990; Sahn and Alderman 1997). Few cross-sectional regressions are likely to measure these differences. While it is standard practice to include age variables in the analysis of nutrition or to run regressions by age group, neither of these approaches is entirely effective for exploring whether the effects of a household resource or government program differ by age group.

Another approach to modeling the dynamics of nutrition using lagged dependent variables is taken by

Bhargava (1995, 1997). Bhargava's studies of adult and child nutrition aim to determine the effects of malnutrition on worker productivity and cognitive development, respectively. These two studies, along with Foster (1995), use data collected as part of multiround longitudinal surveys that differ appreciably from LSMS surveys. Nevertheless, with the possible exception of credit flows, the variables used in Foster's study are commonly collected in LSMS-type surveys. Bhargava's studies use information on individual food intake and, in one of the studies, psychological tests.

Occasionally it is desirable to model how past growth affects current growth—say, to test the hypothesis of catch-up growth or model the impact of short-term economic shocks. Since it takes two observations on the stock of height or weight to indicate the flow or change, questions about dynamics of growth may require a panel that provides at least three observations. LSMS surveys have occasionally carried out repeat observations of the same panel of households but have rarely provided more than two observations from each household. Thus only a few existing LSMS surveys are well suited to address issues of child growth or sequential consequences of health shocks.

Even when levels of nutrition, rather than growth, are being studied, having panel data allows analysts to address errors in measurement. It also allows analysts to distinguish the effects of nutrition from the effects of other human capital investments. Because household health and schooling investments are determined by the same allocation processes that the household applies to the use of its resources, in models of the demand for schooling conditional on nutrition it is a challenge to distinguish the effects of health from the effects of other household characteristics. One way of approaching the problem is to use prices in an instrumental variables method. However, all current prices arguably belong in a household's budget constraint, and thus are not useful as instrumental variables in this context. As such, it is not possible in a cross-section to use current prices to identify nutrition in a simultaneous model of schooling decisions; this may be possible in panel data sets. (Glewwe and Jacoby 1995 use maternal height as an identifying variable.)

Taking height as a stock variable, one alternative to cross-sectional analysis is to use past shocks to identify the impact of nutrition on current conditions. This approach depends on the availability of panel data that include repeated measurements of nutritional status

and prices, varying over time and space.<sup>9</sup> This above approach must also assume, consistent with most evidence, that catch-up growth is imperfect. While there is limited evidence on the difference between using current prices as opposed to lagged prices to identify nutrition in models of current decisionmaking, Alderman and others (1997) provide an example for which the bias is shown to be appreciable.

### Summary

There are two main advantages to anthropometric data gathered in national, multitopic household surveys like the LSMS. The first advantage is the broad range of analysis that can be done with these data because data on many other topics are gathered in the same survey at the same time. The second advantage is that these data come from a nationally representative sample. In contrast, clinic-based anthropometric data are biased because the children who attend clinics are not a random sample of the population. Moreover, the extent and direction of the bias in clinic data is not generally known (Grosh, Fox, and Jackson 1991).

There are also limitations to the anthropometric data gathered from LSMS-type surveys. For example, it is not practical to take blood samples to measure micronutrient levels in this kind of survey. And the relatively small samples (2,000–5,000 households) used by these surveys make it inappropriate to disaggregate malnutrition rates to local levels such as the province or district—limiting the extent to which these data can be used to prioritize regions or districts for programs to reduce malnutrition.

Box 10.1 lists the policy issues that can and cannot be analyzed with the anthropometric data from LSMS-type surveys.

### The Draft Anthropometry Module

As is virtually always the case, the choice of which variables to collect in a multitopic survey will be determined by country-specific factors such as what information already exists, what policy debates are current, and what resources are available. However, it is possible to make a few general suggestions based on previous analytical work carried out using LSMS survey data:

- Since the range of analysis that can be undertaken with any data set exceeds the imagination of any

#### Box 10.1 Policy Issues and Anthropometric Data

*Issues that can be analyzed using LSMS survey data*

- Monitoring the nonincome dimensions of poverty.
- Assessing the nutritional status of the population over time and space.
- Indicating and analyzing the distribution of resources by age and gender within a household.
- Providing perspective on targeting errors.
- Gauging the role of incomes and prices in nutritional status.
- Gauging the role of education in nutritional status.
- Estimating the labor productivity returns to nutritional investments.

*Issues that can be analyzed using LSMS surveys with specialized modules (including a community questionnaire)*

- Demonstrating the impact of the quality of health and sanitation infrastructure.
- Measuring the degree to which private and public resources (including education) complement each other in raising the nutritional status of the population.

*Issues that require analysis of additional non-LSMS data or special samples*

- Estimating the educational returns to nutritional investments.
- Conducting project evaluations.
- Distinguishing between short-term and long-term responses to income, price, and health shocks.
- Identifying credit constraints.

*Issues that cannot be analyzed using household survey data*

- Providing the necessary information to target interventions at community or district levels.

*Note:* To paraphrase Griliches (1984), imperfect data has the virtue of allowing a researcher to show his or her creativity. While this box implies that LSMS surveys are relatively unsuited for shedding light on some issues, a researcher can often transcend these limitations with novel solutions. Some of the illustrations given in this chapter are exceptions to the classifications in this table.

single researcher, it is desirable to collect both weights and heights, even though the conceptual advantages of having both measures have not been fully realized in existing research.

- To promote the use of weight and height measures as indicators of community health, these measures should be collected for all children under the age of 60 months and preferably also for older children. The descriptive statistics that use this information should be disaggregated by age and gender.

- Adult height and weight should be measured, at least in full-length multitopic surveys. Without data on parents (particularly mothers), regression analyses explaining height and weight indicators for children are likely to be seriously biased. Collecting data on parental height and weight also provides useful insight into household decisionmaking.
- Any module studying adult health should measure weights as well as heights. However, since BMI can be very sensitive to intensity of labor, fertility choices, and sample selection, analysis of adult weights will need to consider these factors.
- If it is to be possible to analyze the determinants of both height for age and weight for height and to use instrumented anthropometric variables in analysis, the range of price and infrastructure variables currently collected by LSMS surveys will have to be expanded. One possibility would be to add modules on health delivery services or community programs to the community or facility questionnaires, but two other approaches could also be considered. First, in some countries it may be possible to make use of GIS and other regional administrative databases as well as LSMS data. (For communities living at high altitudes, the GIS data should include altitude information.) Second, a search for price series data can be undertaken concurrent with the preparation of the multitopic survey. Neither type of information is likely to correspond precisely to the survey's sample clusters, but it may be possible to determine which data collection points for prices should be associated with which sample clusters.
- It is better to use a relatively uncommon measuring mechanism, such as mid-upper arm circumference, than not to collect anthropometric measurements at all. The information that can be obtained with that technique, while not perfect, provides analysts with a sound basis from which to draw conclusions about child welfare. It is possible that the main reason why the mid-upper arm circumference measurement is unpopular with survey analysts is that they are not used to using it. If some future LSMS-type surveys take arm circumference measurements in addition to weight and height measurements, this will provide experiential evidence on the use of this measurement that can be taken into account in designing subsequent surveys.

Box 10.2 provides cautionary advice about the extent to which this module has been used in previous LSMS

#### Box 10.2 Cautionary Advice

- *How much of the draft module is new and unproven?* None. Data on the height, weight, and age of children have been gathered in LSMS and other surveys in many countries. It has been less common for surveys to collect these data for adults. Nevertheless, this has been done in enough countries that it is now a straightforward process. Mid-upper arm circumference has been measured in LSMS surveys only in Vietnam, but no difficulties were encountered in that country.
- *How well has the module worked in the past?* Where high-quality data have been gathered, they have proved valuable for analysis. However, collecting anthropometric data is a painstaking exercise with logistical implications that must be carefully examined. Quality control, supervision, and extensive training for interviewers are essential. Scales and measuring boards must be procured (a process that can sometimes take a surprisingly long time) and these items must be carried around by the team members responsible for anthropometry. (Team members often complain about carrying the measuring boards, since these boards can be heavy and awkward.) Even after logistics have been taken care of, in some surveys the resulting data have not been of adequate quality. The final section of this chapter discusses how to maintain quality of anthropometric data; United Nations (1986) is the standard handbook on the topic.
- *Which parts of the module most need to be customized?* The module is more standard than most in terms of what items of data to collect for each person. The main decision is whether to measure all household members or only young children. This decision should take into account both the analytical objectives of the survey and any logistical constraints. If the survey is done in a country where many people do not know how old they are, it may be necessary to devise a local calendar of important events and their dates to help accurately identify the ages of respondents.

surveys, and the extent to which it has provided useful data in these surveys.

#### Notes on the Anthropometry Module

This section addresses issues of how the LSMS survey questionnaire should be applied in the field and how descriptive data from the module should be presented.

#### Measurement Issues

Although anthropometric measurements are believed to be unbiased, objective indicators of the health of a

population, in fact these measures are by no means error-free. Kostermans (1994) illustrates that even random errors in measurements can lead to biased estimates of malnutrition rates. The most common problems at the interview stage include: a household or community setting in which individuals are unable or unwilling to remove their clothes when they are weighed; the infrequent recalibration of scales; a tendency for interviewers to round weights up or down to the nearest kilo or half-kilo (less of a problem for height measurements since they are taken in centimeters, a unit that comprises a small portion of total height); and a tendency for interviewers to under-record the lengths of children under the age of two because the children are not extended to their full length while being measured.

In addition, errors are often introduced by imperfect age measures. Ages are frequently rounded up or down to the nearest year or half-year. And in some cases interviewers may record children as above the age cutoff for measurement—if there is such a cutoff—in order to reduce their own workload. This is less of a problem where birth or health records are available, as these provide a fairly reliable measure of age even if the registry of birth is delayed and somewhat inaccurate. This error declines as a percentage of actual age because the time since the original entry on the health record is usually known with certainty by the parent. In a longitudinal study the time interval since first measurement is subject only to error in an interviewer's recording of dates, and not to errors in parents' memories of birth dates.

Other than improving the training of interviewers or going back to households to check outlier measurements, there are no easy solutions to these problems. The United Nations' suggestion to improve age recall by constructing community events calendars to temporally situate births (United Nations 1986) is of limited use in a national sample as it is unlikely that such calendars will be equally pertinent for all households in a national sample. However, age recording can be improved by recording data in months rather than in years and months (which often leads to inconsistency in units) and by including a question on an individual's date of birth in order to double-check his or her current age. When the age determined from the community events calendar conflicts with the age determined from the given birth date (and clinic records are unavailable), a decision rule, based on inde-

pendent validation of similar approaches, is needed. The author is not familiar with any examples of such a rule.

Kostermans (1994) and Gibson (1990) list the sources of the most common measurement errors, most of which could be solved by improving the training of interviewers. (An extensive discussion of field techniques can be found in United Nations 1986.) LSMS survey teams often include staff who specialize in anthropometric measurements; these people can be given intensive training. Another way to reduce error is for interviewers to repeat their measurements, taking the second measurement either immediately after the first measurement or at a later date. Repeating the collection of height and weight measurements (usually after two weeks) need not be very costly if the repeat visit coincides with the fielding of the expenditure module. The costs of repeat visits may also be minimized by repeating only a subset of the initial visits or revisiting only households whose observations were questionable. Nevertheless, repeating these measurements may yield fewer gains than might be anticipated. If, for example, interviewers repeat the same measurement mistakes they made the first time, or if they record the initial measurement a second time without actually taking a second measurement, the error will not be corrected.

An additional approach, inherent in LSMS surveys but seldom used in analysis, is to use redundant information as instrumental variables. Repeated measurements of height can be used as instrumental variables when height is a right-hand-side variable. Even in the absence of repeated measurements, heights can be instrumented for weights or mid-upper arm circumference. However, age standardization presents a problem, as age error will be common across measurements. And weight for height is not an appropriate instrument for height for age, as the two are not expected to be correlated. Worse, the two may be correlated due to a common measurement error in height.

#### *The Presentation of Malnutrition Rates*

Given that heights and weights of children are highly age-dependent, it is seldom useful to study anthropometry means and distribution of height and weight measures that have not been standardized for age. Summary statistics are most commonly presented in terms of an international reference—which is includ-

ed in the commonly used nonproprietary software produced and distributed by the Center for Disease Control in Atlanta and by the WHO.<sup>10</sup>

Although WHO (1995) advocates using a single international reference, the use of references based on a population of American children is often criticized as an inappropriate benchmark for assessing the health of children in developing countries. This criticism is not justified. It is often found that children from privileged or middle-class families in developing countries have height and weight distributions that do not differ from international references (WHO 1995; Habicht and others 1974; Graitcher and Gentry 1981). Also, the heights and weights of refugee children from Asian countries have been shown to converge rapidly with those of children in United States as the refugee families face fewer economic constraints (Yip, Scanlon, and Trowbridge 1992). Moreover, the use of international references is consistent with the fact that across many countries there is a relationship between mortality risk and malnutrition as measured using such references (Pelletier 1994).<sup>11</sup>

Height data are often reported in terms of Z scores for height for age, derived from the unit normal curve after subtracting the age- and gender-specific means of height from the observation and after dividing by the corresponding standard deviation. Malnutrition rates can then be presented in terms of the percentage of the population that falls below  $-2$  Z scores. In the reference population 2.3 percent have Z scores below this level, while 16.0 percent are below  $-1$  Z score. These levels might be expected for a normal population, and provide a basis for comparison. However, there is no sharp difference in risk of mortality or functional impairment at this or any other commonly used cutoff level. Thus there is a need to determine and present the distribution of anthropometric status and not just malnutrition rates.

Malnutrition rates are often reported in terms of the number of children below a stated percentile of the reference distribution or the number whose heights are less than a certain percentage of the median height. The second way of reporting is not as useful as the first since it does not allow a comparison to be made with an expected distribution. For example, a point at 90 percent of the reference height will represent a different percentile of the reference population depending on the child's age (Waterlow and others 1977). The difference in reported malnutrition can

be substantial; data from the 1987 Ghana Living Standards Study show that the rate of malnutrition among rural (urban) children was 34.8 (22.0) using  $-2$  Z scores, while it was 22.8 (12.3) using heights below 90 percent of the reference median (Alderman 1990).

Because low as well as high values of BMI have health consequences, summary statistics are often presented in terms of ranges. For example, James, Ferro-Luzzi, and Waterlow (1988) suggest that BMI levels between 18.5 and 23.0 should be considered normal, with values above 23.0 indicating that an individual is over-weight. Individuals falling below 18.5 are assigned to one of three categories of energy deficiency ranging from grade I (BMI between 18.5 and 17.0) to grade III (BMI below 16.0). There is some indication that the distribution in a population may vary by gender; in Ghana, women were more likely than men to be classified as undernourished and as overweight (Alderman 1990). This result, in conjunction with the fact that Waaler's data on mortality do not show increased mortality for individuals with BMI around 23.0, suggests that 23.0 is a low cutoff point for defining obesity.

## Notes

The author would like to express gratitude to Nauman Illias for the assistance provided in preparing this chapter. Jere Haas and David Pelletier provided valuable advice on technical points. The author also thanks John Strauss, Duncan Thomas, and Alfred Zerfas for comments on an earlier draft.

1. Excellent discussions of a broad range of issues regarding the anthropometric assessment of individuals and populations can be found in WHO (1995) and Gibson (1990). A review of the ways anthropometry has been used in monitoring or promoting the growth of children can be found in Ruel (1995).

2. In Africa the nutritional status of girls is as good as or better than that of boys (Svedberg 1990), while in Asia girls often appear to be disadvantaged. Yet even within Asia the evidence is by no means uniform (Harriss 1995). Nutritional data from the Pakistan Integrated Household Survey and most other studies from that country do not indicate that girls have higher rates of malnutrition than boys (Alderman and Garcia 1994), even though the evidence is clear that girls have higher mortality rates than boys. Globally, there is some evidence that girls are more resistant to stress than boys (Yip 1996). Overall, equal levels of malnutrition in environments with periodic food shortages and high morbidity patterns may reflect the fact that households allocate more of their resources to boys than to girls. However, this conjecture has not yet been subjected to rigorous testing.

3. The problems associated with evaluating programs using regular LSMS samples are that few programs are sufficiently widespread for beneficiaries to be represented in sufficient numbers within an LSMS sample and that it is hard to control for site and individual selection using cross-sectional data. It is often necessary to introduce an element of experimental design into the survey to address these issues. See Newman, Gertler, and Rawlings (1994).

4. Examples of studies that have used LSMS data to model the determinants of nutrition include Alderman (1990), Glewwe (1995), Piwoz (1995), Sahn (1992), and Strauss (1990). For a survey of the wider literature see Strauss and Thomas (1995).

5. A further distinction is made between being small and becoming small (Beaton 1989). Becoming small is an indication of stress, which affects health and is of particular importance in a clinical setting where repeated measurements are commonly used, for example, in growth promotion programs.

6. Rather than exclude children with missing parents, dummy variables are included for missing parents. The coefficients of such variables should equal roughly the product of average height and the coefficient of parental height. However, if the absence of a parent has a direct impact on child health due to changes in caregiving or an indirect impact through sample selection, this calculation may become less accurate.

7. If parental mid-upper arm circumference is included, the coefficient of income [in Vietnam?] drops to 0.104 ( $t=1.833$ ). The coefficients of both maternal and paternal arm circumference have  $t$  scores above 6, while only maternal height is significant in the analogous regression.

8. For a wider discussion of the range of analyses that use anthropometry modules from LSMS surveys in conjunction with other data, see Strauss and Thomas (1995).

9. Current observations on nutrition could plausibly be taken as the function of current and lagged prices. Thus, in principle, a cross-sectional data set could be linked with a time series on prices. However, this has not yet been explored.

10. Kostermans (1994) indicates that further information can be acquired from the Division of Nutrition, The Center for Disease Control, 2600 Clifton Rd., MS A08, Atlanta, GA 30333, and from the Nutrition Unit, WHO, 1211 Geneva 27, Switzerland.

11. WHO (1995) recommends that the reference data be updated to correct some technical drawbacks including a discontinuity in the standards at the age of 2 (Dibley and others 1987). However, this recommendation makes it quite clear that the organization still endorses a single international standard.

## References

Adair, Linda, and Barry Popkin. 1988. "Birthweight Maturity and Proportionality in Filipino Infants." *Human Biology* 60 (2): 319-40.

Alderman, Harold. 1990. "Nutritional Status in Ghana and its Determinants." Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper 3. World Bank, Washington, D.C.

———. 1993. "New Research on Poverty and Malnutrition: What Are the Implications for Policy?" In Michael Lipton and Jacques van der Gaag, eds., *Including the Poor*. Washington, D.C.: World Bank.

———. 1995. "Information as an Input into Food and Nutrition Policy Formation." In Per Pinstrup-Andersen, David Pelletier, and Harold Alderman, eds., *Enhancing Child Growth and Nutrition in Developing Countries: Priorities for Action*. Ithaca, N.Y.: Cornell University Press.

Alderman, Harold, and Marito Garcia. 1994. "Food Security and Health Security: Explaining the Levels of Nutritional Status in Pakistan." *Economic Development and Cultural Change* 42 (3): 485-507.

Alderman, Harold, Jere Behrman, Victor Lavy, and Rekha Menon. 1997. "Child Nutrition, Child Health, and School Enrollment: A Longitudinal Analysis." Policy Research Working Paper 1700. World Bank, Poverty and Human Resources Division, Policy Research Department, Washington, D.C.

Anand, Sudhir, and Martin Ravallion. 1993. "Human Development in Poor Countries: On the Role of Private Income and Public Services." *Journal of Economic Perspectives* 7 (Winter): 133-50.

Beaton, George. 1989. "Small but Healthy? Are We Asking the Right Questions?" *Human Organization* 48 (1): 11-15.

Behrman, Jere R. 1996. "Impact of Health and Nutrition on Education." *World Bank Research Observer* 11 (February): 23-37.

Bhargava, Alok. 1995. "Econometric Analysis of Psychometric Data: A Model for Kenyan Schoolers." University of Houston, Department of Economics, Houston, Tex.

———. 1997. "Nutritional Status and the Allocation of Time in Rwandese Households." *Journal of Econometrics* 77 (1): 277-95.

Bouis, H., and L. Haddad. 1992. "Are Estimates of Calorie-Income Elasticities Too High? A Recalibration of the Plausible Range." *Journal of Development Economics* 39 (2): 333-64.

Briscoe, J., J. Akin, and D. Guilkey. 1990. "People Are Not Passive Acceptors of Threats to Health: Endogeneity and Its Consequences." *International Journal of Epidemiology* 19 (1): 147-53.

Calle, Eugenia, Michael Thun, Jennifer Petrelli, Carmen Rodrigues, and Clark Heath. 1999. "Body-mass Index and Mortality in a Prospective Cohort of U.S. Adults." *The New England Journal of Medicine* 341 (15): 1097-1105.

Dibley, Michael, Norman Staehling, Phillip Nieburg, and Frederick Trowbridge. 1987. "Interpretation of Z-Score Anthropometric Indicators Derived from the International Growth References." *American Journal of Clinical Nutrition* 46 (5): 749-62.

- Filmer, Deon. 1995. "The Intrahousehold Allocation of Health and Cognitive Skills in Developing Countries." Ph.D. diss. Brown University, Department of Economics, Providence, R.I.
- Fogel, Robert. 1994. "Economic Growth, Population Theory, and Physiology: the Bearing of Long-Term Processes on the Making of Economic Policy." *American Economic Review* 84 (3): 369–95.
- Foster, Andrew. 1995. "Prices, Credit, Markets and Child Growth in Low-Income Rural Areas." *Economic Journal* 105 (May): 551–70.
- Foster, Andrew, and Mark Rosenzweig. 1994. "A Test for Moral Hazard in the Labor Market: Contractual Arrangements, Effort, and Health." *Review of Economics and Statistics* 76 (2): 213–27.
- Gertler, Paul, Paul Glewwe, and Ninez Ponce. 1998. "Poverty, Growth and Nutrition." In David Dollar, Paul Glewwe, and Jennie Litvak, eds., *Household Welfare and Vietnam's Transition to a Market Economy*. Washington, D.C.: World Bank.
- Gibson, Rosalind. 1990. *Principles of Nutritional Assessment*. Oxford: Oxford University Press.
- Glewwe, Paul. 1999. "Why Does Mother's Schooling Raise Child Health in Developing Countries? Evidence from Morocco." *Journal of Human Resources* 34 (1): 124–59.
- Glewwe, Paul, and H. Jacoby. 1995. "An Economic Analysis of Delayed Primary School Enrollment in a Low-income Country: The Role of Early Childhood Malnutrition." *Review of Economics and Statistics* 77 (1): 156–69.
- Glewwe, Paul, and J. van der Gaag. 1990. "Identifying the Poor in Developing Counties: Do Different Definitions Matter?" *World Development* 18 (6): 803–14.
- Golden, Michael. 1994. "Is Complete Catch-Up Possible for Stunted Malnourished Children?" *European Journal of Clinical Nutrition* 48 (1): S58–S70.
- Graitcher, P. L. and E. M. Gentry. 1981. "Measuring Children: One Reference for All." *Lancet* (2): 297–99.
- Griliches, Zvi. 1984. "Econometric Data Issues." In Z. Griliches and M. Intriligator, eds., *Handbook of Econometrics*. Vol. 3. Amsterdam: North Holland Press.
- Grosh, Margaret, Kristin Fox, and Maria Jackson. 1991. "An Observation on the Bias in Clinic-Based Estimates of Malnutrition Rates." Policy Research Working Paper 649. World Bank, Washington, D.C.
- Habicht, J-P., R. Martorell, C. Yarbrough, R. Malina, and R. Klein. 1974. "Height and Weight Standards for Preschool Children: How Relevant are Ethnic Differences in Growth Potential?" *Lancet* 1 (7,858): 611–14.
- Haddad, Lawrence, and Howarth Bouis. 1991. "The Impact of Nutritional Status on Agricultural Productivity: Wages Evidence from the Philippines." *Oxford Bulletin of Economics and Statistics* 53 (1): 45–68.
- Harriss, Barbara. 1995. "The Intrafamily Distribution of Hunger in South Asia." In Jean Dreze, Amartaya Sen, and Artar Hussain, eds., *The Political Economy of Hunger. Selected Essays*. Oxford: Clarendon Press.
- Higgins, Paul, and Harold Alderman. 1997. "Labor and Women's Nutrition. A Study of Energy Expenditure, Fertility, and Nutritional Status in Ghana." *Journal of Human Resources* 32 (3): 577–95.
- Horton, Susan. 1988. "Birth Order and Child Nutritional Status: Evidence from the Philippines." *Economic Development and Cultural Change* 36 (2): 341–54.
- Kostermans, Kees. 1994. *Assessing the Quality of Anthropometric Data: Background and Illustrated Guidelines for Survey Managers*. Living Standards Measurement Study Working Paper 101. Washington, D.C.: World Bank.
- Lavy, Victor, John Strauss, Duncan Thomas, and Phillippe de Vreyer. 1996. "Quality of Health Care, Survival and Health Outcomes in Ghana." *Journal of Health Economics* 15 (3): 333–57.
- Lutter, Chessa, Jose Mora, Jean-Pierre Habicht, Kathleen Rasmussen, Douglas Robson, and Guillermo Herrera. 1990. "Age-Specific Responsiveness of Weight and Length to Nutritional Supplementation." *American Journal of Clinical Nutrition* 51 (3): 359–64.
- James, W.P.T., A. Ferro-Luzzi, and J.C. Waterlow. 1988. "Definition of Chronic Energy Deficiency in Adults." *European Journal of Clinical Nutrition* 43 (12): 969–81.
- Martorell, R., J. Rivera, H. Kaplowitz, and E. Pollitt. 1992. "Long-term Consequence of Growth Retardation During Early Childhood." In M. Hernandez and J. Argente, eds., *Human Growth: Basic and Clinical Aspects*. Amsterdam: Elsevier Science Publishers.
- Newman, John, Paul Gertler, and Laura Rawlings. 1994. "Using Randomized Central Designs in Evaluating Social Sector Programs in Developing Countries." *World Bank Research Observer* 9 (2): 181–201.
- Pelletier, David. 1994. "The Potentiating Effects of Malnutrition in Child Mortality: Epidemiologic Evidence and Policy Implications." *Nutrition Reviews* 52 (12): 409–15.
- Pitt, M., M. Rosenzweig, and M. D. Hassan. 1990. "Productivity, Health, and Inequality in the Intrahousehold Distribution of Food in Low-Income Countries." *American Economic Review* 80 (5): 1139–56.
- Piwoz, Ellen. 1995. "Undernutrition in Nicaraguan Preschool Aged Children: Prevalence, Determinants, and Policy Implications." Latin American and Caribbean Region Economic Note 1. World Bank, Washington, D.C.
- Pollitt, Ernesto. 1990. *Malnutrition and Infection in the Classroom*. Paris: United Nations Educational, Scientific and Cultural Organization.
- Ravallion, Martin. 1993. *Poverty Comparisons: Fundamentals of Pure and Applied Economics*. Vol 56. Chur Switzerland: Harwood Academic Press.

- Rosenzweig, Mark R. and T. Paul Schultz. 1988. "The Stability of Household Production Technology: A Replication." *Journal of Human Resources* 23 (4): 535–49.
- Ruel, M. 1995. "Growth Monitoring as an Educational Tool, an Integrating Strategy, and a Source of Information: A Review of Experience." In P. Pinstrip-Andersen, D. Pelletier, and H. Alderman, eds., *Enhancing Child Growth and Nutrition in Developing Countries: Priorities for Action*. Ithaca, N.Y.: Cornell University Press.
- Sahn, David E. 1994. "The Contribution of Income to Improved Nutrition in Cote d'Ivoire." *Journal of African Economies* 3 (1): 29–61.
- Sahn, David, and Harold Alderman. 1997. "On the Determinants of Nutrition in Mozambique: The Importance of Age-Specific Effects." *World Development* 25 (4): 577–88.
- Steckel, Richard. 1995. "Stature and the Standard of Living." *Journal of Economic Literature* 33 (4): 1903–40.
- Strauss, John. 1990. "Households, Communities and Preschool Children's Nutrition Outcomes: Evidence from Rural Cote d'Ivoire." *Economic Development and Cultural Change* 38 (2): 231–61.
- Strauss, John, and Duncan Thomas. 1995. "Human Resources: Empirical Modeling of Household and Family Decisions." In Jere Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics*. Vol. 3A Amsterdam: North Holland.
- . 1996. "Measurement and Mismeasurement of Social Indicators." *American Economic Review* 86 (2): 30–34.
- . 1998. "Health Nutrition and Economic Development." *Journal of Economic Literature* 36 (2) June: 766–817.
- Subramanian, S., and A. Deaton. 1996. "The Demand for Food and Calories." *Journal of Political Economics* 104 (February): 133–62.
- Svedberg, Peter. 1990. "Undernutrition in Sub-Saharan Africa: Is There a Gender Bias?" *Journal of Development Studies* 26 (3): 469–86.
- Thomas, Duncan. 1994. "Like Father, Like Son: Like Mother, Like Daughter: Parental Resources and Child Height." *Journal of Human Resources* 29 (4): 950–88.
- Thomas, Duncan, and John Strauss. 1997. "Health and Wages: Evidence on Men and Women in Urban Brazil." *Journal of Econometrics* 77 (1): 159–85.
- Thomas, Duncan, Victor Lavy, and John Strauss. 1996. "Public Policy and Anthropometric Outcomes in Côte d'Ivoire." *Journal of Public Economics* 61 (2): 155–92.
- United Nations. 1986. *How to Weigh and Measure Children: Assessing the Nutritional Status of Children in Household Surveys*. National Household Survey Capability Programme. New York: United Nations.
- UNDP (United Nations Development Programme). 1990. *Human Development Report 1990*. New York: Oxford University Press.
- Vella, Venanzio, A. Tomkins, A. Borghesi, G. Migliori, J. Ndiku, and B. Adriko. 1993. "Anthropometry and Childhood Mortality in Northwest and Southwest Uganda." *American Journal of Public Health* 83 (11): 1616–18.
- Victora, Cesar. 1992. "The Association between Wasting and Stunting: An International Perspective." *Journal of Nutrition* 122 (5): 1105–10.
- Waller, H. 1984. "Height, Weight, and Mortality: The Norwegian Experience." *Acta Medica Scandinavica* 77 (Supplement): 1–56.
- Waterlow, J.C., R. Buzina, W. Keller, J.M. Lane, N.Z. Nichaman, and J.M. Tanner. 1977. "The Presentation and Use of Height and Weight Data for Comparing the Nutritional Status of Groups of Children under the Age of Ten Years." *Bulletin of the WHO* 55 (4): 489–98.
- WHO (World Health Organization). 1995. *Physical Status: The Use and Interpretation of Anthropometry*. Technical Report Series 854. Geneva: World Health Organization.
- Yip, R. 1996. Electronic communication.
- Yip, R., K. Scanlon, and F. Trowbridge. 1992. "Improving Growth Status of Asian Refugee Children in the United States." *Journal of American Medical Association* 267 (7): 937–40.
- Zerfas, Alfred. 1991. "Choice of Nutritional Status Indicators for Young Children in Public Health Programs." Latin America and the Caribbean Technical Department Report 8. World Bank, Washington, D.C.

# 11

## Transfers and Other Nonlabor Income

*Andrew McKay*

This chapter discusses sources of household income that are not covered in Chapters 9, 18, and 19 (on income from wage employment and self-employment in household enterprises or agriculture), in Chapter 12 (on ways to estimate imputed rent from owner-occupied dwellings), or anywhere else in this volume. The income sources covered by this chapter will collectively be referred to as miscellaneous income. Miscellaneous income consists predominantly of current transfers, but it also includes rental income and other income such as interest or lottery winnings.

There are two main reasons for collecting data on transfers and other nonlabor income. One reason is to construct an estimate of total household income when the questionnaire also gathers sufficiently detailed data on housing, agriculture, economic activities, and non-farm enterprises. A second reason is that many of the income sources covered in this chapter are themselves of considerable interest to policymakers and analysts; particularly interesting to policymakers and analysts are households' incomes from public transfers and private interhousehold transfers.

The first section of this chapter sets out some of the main policy issues that can be analyzed with data on transfers and other nonlabor income. The second section discusses the kind of data on these income components that must be collected to analyze the policy issues identified in the first section. The third section introduces standard and short versions of modules designed to collect these data in an LSMS-type multitopic household survey. (The modules are presented in Volume 3.) The fourth and final section provides notes and comments on the proposed modules.

### **Policy Issues**

Most of the income received by households in developing and transition countries comes from the wage employment of their members or from their agricultural or other household enterprises. However, many households also receive income from a variety of other sources, including public transfers, private transfers, and rent.

Households receive rental income if they rent out their assets (land, dwellings, farm equipment) to others.<sup>1</sup> Since only wealthy households usually own sufficient assets to rent these assets to others, wealthy households are the only households likely to receive rental income. Similarly, only well-off households are likely to earn incomes from such sources as interest, dividends, or lottery winnings—incomes which in any case are usually not very substantial.

The vast majority of households receive their most significant miscellaneous income from transfers of various kinds. These may be private transfers from other households (possibly in the same extended family) or government transfers such as state pensions,

unemployment benefits, or family allowances. Households may also receive transfer income from companies, possibly in the form of pensions or dividends, or from nongovernmental organizations (NGOs), which may also provide transfers in the form of food or other goods. Some transfers, especially transfers from other households or from a company, may be sent from outside the country.

Although transfers and other nonlabor income sources are not generally as important as income sources resulting from labor, for some households miscellaneous sources are very important. These households are often of particular interest to policymakers. For example, this may be because many households that rely heavily on miscellaneous income are poor, or receive the vast majority of their income from government transfers, or are economically vulnerable due to their reliance on uncertain private transfers. Thus it can be very important to identify and, if possible, measure such sources—especially in circumstances where policymakers and analysts are interested in exploring the determinants of poverty.

Box 11.1 discusses the importance of transfers and other nonlabor income sources for households in three developing countries. The data in this box suggest that these sources are quite important even for average households (Box Table 11.1), and very important for some households.

Table 11.1 summarizes issues that can be addressed using data on transfers and other nonlabor income. These are discussed in turn below.

#### *Measuring and Analyzing Total Household Income*

Estimates of total household income can be used to address various policy and analytical issues, as discussed in Chapter 17. When survey planners decide they want to estimate total household income, the questionnaire must, as far as reasonably possible, collect data on all of the household's prospective sources of income. To accomplish this it is necessary to comprehensively identify income sources for each household and to measure income from these sources as accurately as possible within a suitable reference period. If the goal is to rank the income sources in terms of the size of their contribution to the household budget (possibly for subgroups of households), the data must be collected at as disaggregated a level as possible. Such disaggregation is also desirable to ensure that, as far as possible, no potential sources of income are missed because they were not explicitly mentioned.

Because some sources of income covered by this chapter—such as transfers from government—are highly sensitive to policy interventions, survey designers may wish to include questions on transfers and other nonlabor income for policy reasons, irrespective of whether the survey aims to measure total household income.

#### *Studying the Incidence and Impact of Public Transfers*

Government transfers can represent a significant proportion of the total income of some households—for example, households in which no member is working. On the other hand, many other households receive no transfer payments from the government at all. Public transfers from the government to households take different forms from one country to another, but some common examples are state pensions, unemployment benefits, food stamps, and disability payments. Some of these transfer payments, such as disability payments, reflect the particular circumstances of the recipient (which may be an individual rather than a household), while other transfers, such as food stamps, may be intended specifically for poor individuals or households.

Two policy questions are of interest regarding government transfers: their incidence—who receives them and in what amounts—and their impact—how receiving government transfers affects individuals' or households' behavior.

**INCIDENCE OF PUBLIC TRANSFERS.** Studying the incidence of public transfers involves examining the relationship between receipt of a given government transfer and the relevant characteristics of recipient households or individuals (such as ages and standards of living). A number of descriptive poverty studies have examined which income groups benefit from public transfers and to what extent these groups benefit. One way the studies have done this is by dividing the population into quintile (or similar) groups according to their standard of living or into poor/not-poor categories, and investigating how receipt of transfers varies among the different groups. This kind of technique has been used in studies of the incidence of food stamp programs (Grosch 1992, 1995a), studies of the extent of targeting error associated with such programs (Cornia and Stewart 1995), and studies of the distributional incidence of social security benefits (Jarvis and Micklewright 1995).

**Box 11.1 The Contribution of Miscellaneous Sources to Household Income**

The degree of detail to which miscellaneous income sources need to be measured depends partly on how much income they contribute to all households and to specific types of households. If such sources contribute little to average household income or to the vast majority of individual households, it may not be necessary to measure the sources accurately or in detail.

The table below explores miscellaneous sources of income as measured in three LSMS surveys: Côte d'Ivoire 1988, Ghana 1988–89, and Peru 1994. (Household income was measured fairly accurately in the Côte d'Ivoire and Peru surveys; in the Ghana survey the estimate of household income is more questionable.) Data are presented for the two categories of miscellaneous income distinguished in the third section of this chapter—income from private interhousehold transfers and “other income” (which includes public transfers, rental income, and income from other sources). Private interhousehold transfers are transfers between households, often consisting of gifts of money or goods from family members who live elsewhere. The precise composition of the “other income” category varies from one country to another, reflecting the different information available from the three different surveys.

While the relative importance of these two miscellaneous income components varies between the countries, taken together they contribute significantly to household income in each country. Miscellaneous income accounts for nearly one-sixth of household income in Ghana and Peru, and just less than one-tenth of household income in Côte d'Ivoire. More importantly, miscellaneous income contributes greatly to the total income of certain households. In these three countries, as in many others, a few households derive all of their (measured) income from private interhousehold transfers or “other income” sources. Each of these countries also contains a significant minority of households for which income from one of these sources accounts for one-quarter or more of total household income.

Such households are liable to be systematically different from households that rely on wage income or self-employment income; some are probably richer than average—such as households that derive most of their income from rent and investments—but others are probably poorer or more vulnerable than average—such as households that rely predominantly on private interhousehold transfers and other transfers to meet their consumption needs. Policymakers may be particularly interested in finding out about the latter group.

**Box Table 11.1 Miscellaneous Sources of Household Income: Three Case Studies**

Share of income	Côte d'Ivoire 1988	Ghana 1988–89	Peru 1994
From private interhousehold transfers			
More than zero	31.4	63.0	38.3
10% or more	14.3	26.7	21.2
25% or more	6.8	16.4	11.8
50% or more	2.7	8.0	4.1
100%	0.4	1.9	0.1
Mean contribution	5.6	12.7	7.7
Sample size	1,556	3,405	3,589
From “other” sources			
More than zero	26.5	27.6	38.0
10% or more	7.2	8.2	18.1
25% or more	4.2	3.6	9.2
50% or more	1.2	1.3	3.8
100%	0.4	0.1	0.3
Mean contribution	2.9	3.1	6.5
Sample size	1,559	3,415	3,568

Source: Author's computations from the data derived from: Côte d'Ivoire Living Standards Survey (1988); Ghana Living Standards Survey Round 2 (1988–89); Encuesta Nacional de Hogares Sobre Medicion de Niveles de Vida, Peru (1994).

**THE IMPACT OF PUBLIC TRANSFERS.** Studying the impact of public transfers requires investigating how households or individuals alter their behavior as a result of receiving public transfers and whether such behavior changes reduce the benefits of these transfers.

One relevant issue is whether and how receiving government transfers affects a household's continued receipt of private interhousehold transfers or the amounts of such transfers. (This issue will be considered in more detail below.) Another important issue is

whether receiving government transfers affects the incentives for household members to seek employment, especially in instances in which these transfers are targeted. Sahn and Alderman (1995) found that a targeted rice ration program in Sri Lanka reduced the work effort of recipient household members by up to two or three days per month. (See Kanbur, Keen, and Tuomala 1995 for some of the underlying theoretical arguments about how targeted transfers affect labor supply.)

A further issue is how households act in response to government programs targeted to particular members of the household. Haddad and Zeller (1997) argued that when a school meals program is targeted toward girls (perhaps because girls have been found to be more malnourished than boys), households may respond by reducing the amount of food girls receive at home. However, the impact of public transfers is not always deleterious. Food stamp programs that require recipient households to use primary health care facilities, and school feeding programs that require recipient households to send their children to school in order to qualify for a transfer have additional benefits for recipients (Grosh 1995b).

The difficulty in modeling all of these issues is working out how recipients of government transfers would have behaved had they not received transfers. When a given survey sample includes both recipients and nonrecipients of government transfers, it becomes possible to include the receipt of government transfers as one of the explanatory variables in a model—say, a model of labor supply. This allows analysts to use the resulting data set to draw preliminary conclusions about the impact of government transfers on the variable in which they are interested (in this case, labor supply). This is essentially the procedure used by Cox and Jimenez (1992, 1995) to model how the receipt of government transfers affects private interhousehold transfers and also by Sahn and Alderman (1995) in their Sri Lanka study.

#### *Studying Private Interhousehold Transfers*

Some households in developing or transition countries receive private transfers of income from other households whose members may or may not be a part of the same family. A common example of a private interhousehold transfer occurs when a young member of a household leaves home to seek work in another location—perhaps in an urban area or abroad—where

this person can earn a higher income and send some of the income back to his or her original household. There are several other circumstances in which private transfers are made, for example, between friends or between neighbors.

In many low-income and lower-middle-income developing countries, private transfers between households are the most important sources of income for a considerable number of households. This is most often the case in countries where social security systems and other mechanisms for distributing public transfers are not well-developed; in such cases these private transfers can provide a vital informal safety net for vulnerable or poor households. (For example, private interhousehold transfers are the most important source of transfers fulfilling this function in Ghana; see Box 11.1.) Therefore, studying the nature and effects of private transfers in poor countries is of considerable analytical and policy interest. The survey questionnaire should collect information on not only households' income from private interhousehold transfers but also households' expenditure on these transfers (Table 11.1). Having data on net flows of these transfers, rather than simply gross flows, ensures that the analyst takes into account the full effects of these transfers on household behavior and welfare.

**PATTERNS OF FLOWS OF PRIVATE INTERHOUSEHOLD TRANSFERS.** Households can be classified into three groups: net recipients of private interhousehold transfers, net donors of such transfers, and households that neither give nor receive such transfers. A useful starting point is to know which households belong to each of these groups and how this status relates to the characteristics of these households—most obviously their standard of living but also factors such as their vulnerability or dependency.<sup>2</sup>

Evidence from a wide number of studies (for example, Bamberger and others 1991, Cox and Jimenez 1995, and Knowles and Anker 1981) confirms the hypothesis that transfers between households generally flow from richer to poorer households, although this obviously needs to be investigated or confirmed in each specific case. However, the available evidence also suggests that large numbers of poor households do not receive such transfers (see Cox and Jimenez 1992, Bamberger and others 1991) and thus lack the informal safety net such transfers may provide.<sup>3</sup> Policymakers should know the limits of the coverage

**Table 11.1 Policy Issues Regarding Nonlabor Income Sources and the Data Needed to Analyze Them**

General issue	Specific area of interest	Data requirements from miscellaneous income modules	Data requirements from other modules	Can this issue be analyzed using data from the standard module?	Can this issue be analyzed using data from the shorter module?
Measurement of total household income		Estimates of household incomes from current transfers, rent, and other miscellaneous sources	Data to estimate all remaining components of household income	Yes	No
Government transfers	Incidence of government transfers	Information on households' receipts of different categories of public transfers	Relevant household characteristics—for example, standards of living and demographics	Yes	Yes
	Impact of government transfers	Information on households' receipts of different categories of public transfers	Data on the variable affected by the transfer (such as labor supply)	Yes	Yes
Private interhousehold transfers	Patterns of flows	Household receipts from private interhousehold transfers	Household expenditure on private interhousehold transfers; measures of such indicators as standard of living, vulnerability, and dependency.	Yes	Yes
	Characterizing flows	Household receipts from private interhousehold transfers, plus information on donors such as their reasons for giving transfers	Household expenditure on private interhousehold transfers; information on characteristics of these transfers; measures of such indicators as standard of living, vulnerability, and dependency	Yes	To some extent
	Modeling determinants	Household receipts from private interhousehold transfers; information on donors such as their reasons for giving transfers; (ideally) information on households' social networks	Characteristics of households such as their standards of living, demographics, and assets	To some extent	To some extent
	Crowding out of public transfers	Household receipts from private interhousehold transfers; information on donors such as their reasons for giving transfers; information on receipts of public transfers	Relevant household characteristics—for example, standards of living and demographics	Yes	To some extent
	Household's ability or willingness to pay for projects	Magnitude of transfers, reasons they were sent, and (ideally) how they change in response to changing economic circumstances		Possibly	Possibly
	Impact on household behavior	Household receipts from private interhousehold transfers; information on donors such as their reasons for giving transfers	Data on variables affected by transfers (for example, labor supply)	Yes	To some extent
	Household sales of capital assets		Households' receipts from the sale of their assets (if not collected elsewhere)	Sales of assets, purchases of assets, and general household characteristics including standard of living	Yes

(Table continues on next page.)

**Table 11.1 Policy Issues Regarding Nonlabor Income Sources and the Data Needed to Analyze Them (continued)**

General issue	Specific area of interest	Data requirements from miscellaneous income modules	Data requirements from other modules	Can this issue be analyzed using data from the standard module?	Can this issue be analyzed using data from the shorter module?
Intrahousehold allocation		Estimates of receipt of government and private transfers by individual household members	Remaining incomes at individual level (where possible), individual characteristics (such as gender and age)	To some extent	No

Source: Author's summary.

of this net and the types of households that it does not cover.

When examining the flow of these transfers, analysts must be able to quantify the magnitudes involved. Are transfer flows sufficient to lift significant numbers of households out of poverty? How much do the transfers alter patterns of income distribution or poverty? To properly study this issue, analysts not only need data on transfers, they also need enough data to estimate total household income—a fact that has implications for the design of the survey questionnaire.

A final concern in collecting data on private interhousehold transfers is the accuracy of information obtained on the existence and magnitude of such transfers. In circumstances where most such transfers do not cross international borders,<sup>4</sup> an initial check can be done by comparing the aggregate receipts of and aggregate expenditures on such transfers to see if they are of similar magnitudes. Where they are not (and assuming that the sample is representative of the population), this suggests that there has been a significant underestimation of one variable or an overestimation of the other. (Even where the magnitudes are similar it is possible that both are underestimated or overestimated to the same extent.)

**CHARACTERIZING FLOWS OF PRIVATE INTERHOUSEHOLD TRANSFERS.** It is highly desirable for analysts to understand the nature of private interhousehold transfers and their role in the household economy. To what extent are such transfers flows between members of the same family living in different households (for example, transfers to elderly parents from their sons or daughters)? Are these transfers mostly one-off payments or are they regular and ongoing? Are these transfers mainly sent for specific purposes (for example, to finance a child's education) or simply to provide general financial support to the recipient household?

A number of studies have investigated some of these issues for particular countries (Knowles and

Anker 1981 for Kenya; Cox and Jimenez 1995 for the Philippines), but the answers to the above questions may differ in each different set of circumstances. Even within any given country, the answers to questions on transfer issues differ from one household to another depending on the characteristics of the household. For some households—especially poor households—private interhousehold transfers function (to varying degrees) as an informal safety net. Such transfers tend to be a more reliable source of income for households that receive them from close family members than they are for households that receive one-off transfers from more distant relatives or from nonrelatives.

**MODELING THE DETERMINANTS OF PRIVATE INTERHOUSEHOLD TRANSFERS.** A descriptive analysis provides a certain amount of information on the pattern of (net) private transfer flows between households. However, more (and more reliable) information can be gained from modeling the determinants of these transfer flows using multivariate techniques, particularly when it is possible to include the factors that influence why certain households receive transfers while others do not. In order to identify these factors, analysts are likely to need information about the social networks to which households belong, as some households may not receive any private transfers because they do not belong to social networks (such as an extended family or a community) that might give them access to such transfers. A household may not expect to receive transfers from relatives in another household because of a family breakup or because nonresident family members are as poor as the household residents.

An analysis of the determinants of private interhousehold transfers was undertaken by Bamberger and others (1991) using data from a survey conducted in a poor district of the city of Cartagena, Colombia. The survey collected information on the social networks to which households belonged. The authors used this and other information to model the determinants of:

- The likelihood of a household belonging to a “socially interactive network,” implying that it had potential access to transfers.
- The likelihood of a household having received a transfer in the previous month.
- Total transfers received by a household.
- Net transfers received by a household.

Each of these factors was related to potentially relevant characteristics of the recipient households (for example, demographic factors, household income, and household assets). The authors undertook a similar analysis of payments of private interhousehold transfers, relating these to the characteristics of donor households.

Many household surveys do not collect information on the social networks to which households belong, making it difficult for analysts to model the determinants of a household’s receipt of transfers. However, a number of authors have attempted to model the determinants of total and net transfers received by a household even without information on social networks. Cox and Jimenez (1995) used data from the Philippines to model the determinants of the net private transfers received by a household as a function of the household’s characteristics. As the model was estimated by ordinary least squares, it is not clear how zero values were treated in this study. (There may be a selectivity bias if the model is based only on nonzero values, and there may be issues of conceptual consistency if zero values are treated in the same way as nonzero values; see Heckman 1990 and Greene 1993, section 22.4.) In their 1992 study of Peru, Cox and Jimenez used a probit model to model the determinants of whether a household received transfers—and also to model the transfer amounts using Heckman’s generalized tobit framework (Heckman 1979). And in a more recent study based on the Cartagena data set, the same authors used an ordered probit model to investigate the factors that influenced whether a household was a net recipient of interhousehold transfers, a net donor of interhousehold transfers, or neither (Cox and Jimenez 1998).

By modeling the determinants of both payments and receipts of private transfers between households, it may be possible to better understand how these transfers change in response to a shock in the economy. An important issue in such circumstances is whether private interhousehold transfers cushion households from the effects of shocks or whether the shocks make pri-

ivate transfer flows less likely or smaller. If a shock has an adverse effect on a household that receives private transfers, this may prompt the households that donate the transfer to increase the amount or frequency of transfers, thus cushioning the effect of the shock on the recipient household. Alternatively, when the household that donates a transfer suffers an adverse shock, it may not be able to continue the transfer—thus adversely affecting the household that had previously benefited from the transfer.

Consider the case of an urban-based household that derives much of its income from wage employment in the formal sector and that makes regular transfers to relatives who live in a poorer household in a rural area. If a contraction in the urban formal sector caused one of the members of the urban household to become unemployed, the income of the household would decrease. This could reduce the magnitude of the transfer that the urban household sent to the rural household. Thus the impact of the contraction in the urban formal sector would be felt not just in an urban area but also in a rural area. In summary, flows of private interhousehold transfers may respond to changes in the circumstances of either a household that makes a transfer or the household that receives this transfer. If these changes result from policy decisions, their effects on transfers should be considered when the effects of policies are being evaluated.

#### **PRIVATE INTERHOUSEHOLD TRANSFERS AND THE EFFECTIVENESS OF REDISTRIBUTIVE PUBLIC TRANSFERS.**

In situations where public and private transfers are close substitutes for each other, increasing the public transfers received by a given group may result in this group receiving fewer private transfers. Households that had given transfers to this group may feel that it is less imperative to do so. Households that had received private transfers may put less effort into maintaining and developing their social networks. Whatever the explanation, if public transfers do “crowd out” private transfers in this way, the distributional incidence of the public transfers (allowing for indirect as well as direct effects) will be different from what it initially appears to be. Assuming that the donors of private transfers are better off than the recipients, the benefits of public transfers to poor households may at least partially accrue to richer households.

This issue has been examined in a series of studies by Cox and Jimenez (1990, 1992, and 1995). In the 1992 and 1995 studies the authors used multivariate

regression techniques to model the determinants of the amount of net private transfer income received by a household as a function of the relevant characteristics of that household. Among these relevant characteristics are the type and magnitude of any public transfers received. If public transfers are crowding out private transfers, the terms representing the receipt of income from public transfers would be expected to have statistically significant negative coefficients; the coefficients of the regression could be used to estimate the response of private transfers to specific changes in public transfers. In the case of Peru (Cox and Jimenez 1992) the authors found that private transfer payments from the young to the old would have been nearly 20 percent higher in the absence of social security pension payments—which suggests that some crowding out was occurring. The authors' study of the Philippines (Cox and Jimenez 1995) suggested that private transfers were strongly responsive to the recipients' income levels. Thus, if unemployment insurance were to be introduced, there would be a significant—though not complete—crowding out of private transfers.

**USING INFORMATION ON PRIVATE TRANSFERS TO ASSESS INDIVIDUALS' ABILITY AND WILLINGNESS TO PAY FOR DEVELOPMENT PROJECTS.** One aim of the study in Cartagena, Colombia (Bamberger and others 1991) was to assess how private transfers could be expected to respond to an urban development project in the area of the city where the survey was conducted. Private transfers were clearly an important source of income for households in this area. The authors argued that when the local community perceives a project of this type as a desirable investment, private transfer income to households in the vicinity of the project can increase substantially. In fact this was not observed in Cartagena—a result the authors attributed to the perception of locals that the project did not represent a good investment for the community. But for positive evidence to support their argument, the authors cited cases from El Salvador and the Philippines (Bamberger and others 1991, chapters 6 and 7). In cases where it seems likely that private transfer income will increase when a development project is introduced, policymakers should bear this factor in mind in assessing the ability and willingness of the local community to pay for the project.

**STUDYING THE IMPACT ON HOUSEHOLDS OF RECEIVING PRIVATE TRANSFERS.** Like public transfers, private

transfers can have an effect on recipient households' behavior. These changes can take a number of different forms that may or may not have been intended by the donor. Unfortunately, it is very difficult and complex to model the way in which receiving private transfers affects the behavior of recipients, not least because the magnitude and even the existence of a private transfer flow might be endogenous (in other words, within the control of the recipient household). This is much less of an issue in modeling the impact of public transfers because the amounts of public transfers and the criteria for their provision are more clearly predetermined.

One example of the complexity of this process is how problematic it is to model the effects of private transfers on the participation of members of the recipient household in the labor market. It is conceivable that the donor would withdraw or reduce the transfer if members of the recipient household stopped working or stopped looking for work as a result of receiving the transfers. Thus both the receipt of private transfer income and the amount received are probably linked to the behavior of the recipient household's members with regard to the labor market. In principle, therefore, the model should make both labor supply and the receipt of private transfers endogenous.

However, some effects of receiving private inter-household transfers are easier to model. While controlling for household income and other relevant factors, Bamberger and others (1991) use multivariate regression analysis to examine how receiving private transfer income affects a household's expenditure on basic and nonbasic needs. In this case, the possible endogeneity of the transfer income as an explanatory variable may be less important, in that it may be difficult for the donor to observe how the transfer income is being spent.<sup>5</sup>

#### *Studying the Sale of Assets by Households*

Revenue from the sale of household assets should not be regarded as income because these sales are capital transactions. Nevertheless, the nonlabor income module in a multitopic household survey may be a good place to collect information on such asset sales to the extent that it is not collected elsewhere in the questionnaire.<sup>6</sup>

Sometimes households sell off their assets simply to buy other assets. Thus selling assets does not necessarily imply a drop in the household's long-term sustainable standard of living; the reverse may in fact be true. On

the other hand, households sometimes sell their assets to finance consumption, reducing both the net worth of the household and its long-term sustainable standard of living. Gathering information on asset sales can help analysts identify and study those households that sell their assets to stabilize their consumption levels in response to either a fall in income or a large unexpected expenditure.<sup>7</sup> As discussed by Corbett (1988), this strategy is usually a last-resort response to a famine or other emergency situation, especially when the assets sold are productive. A household is clearly in a desperate situation if it is engaged in selling its assets merely to finance consumption. Thus, if a survey questionnaire collects data on the reasons why households sell their assets, this will enable analysts to identify households that may be in danger of falling into poverty.

#### *Studying the Impact of Transfers and Other Nonlabor Income on Intrahousehold Allocations*

When users of survey data wish to study intrahousehold dynamics, it is desirable to collect information on transfers and other nonlabor income sources at the level of the individual rather than the entire household. For example, because women tend to ensure a household's food security and improve household members' nutritional outcomes (Quisumbing and others 1995; Haddad and Zeller 1997), it is likely to be a good idea to ensure that transfers (particularly government transfers) are given directly to the women in the household. The issues involved in studying such intrahousehold dynamics are discussed in Chapter 17 on measuring total household income.

#### *Other Issues*

In some countries or situations analysts may be interested in looking at the distributional incidence of other miscellaneous income sources—transfers received from NGOs or, in Islamic countries, transfers of *zakat* (a form of Islamic charity intended to benefit the poor). The principles involved in performing such an analysis are similar to those discussed above, although their relevance to policy may be less clear-cut. Some of the remaining nonlabor income categories—such as interest or dividends earnings (which will almost always disproportionately benefit richer households) or lottery winnings—may need to be measured only to construct estimates of total household income, since otherwise they are unlikely to be of much relevance to policymakers or analysts.

### **Data Requirements**

The planners of any multitopic household survey should design the survey questionnaire to collect the data needed to analyze the policy issues that they have identified as most relevant to the country studied. If the planners have decided to measure total household income, the data collected on current transfers, rental income, and other miscellaneous income should be as comprehensive and accurate as is reasonably possible. Beyond this, whatever data are collected on these income sources will depend on exactly what the likely policy and analytical applications of the data will be. In most circumstances government transfers and private interhousehold transfers are the income sources of greatest interest to policymakers and analysts. In some cases revenues from the sale of capital assets are also likely to be of interest.

This section of the chapter discusses some of the main issues to be addressed in designing modules that will collect sufficiently reliable data on these transfer and nonlabor income components. The implications for data collection of the various analytical and policy issues discussed in the first section of this chapter are set out in Table 11.1.

#### *Does the Survey Aim to Collect the Data Necessary for Measuring Household Income?*

The issue of whether an LSMS survey should collect the data necessary to measure total income is discussed in depth in Chapter 17. In that chapter it is argued that a standard-length multitopic household survey should aim to measure total household income, whereas a shorter survey should not attempt to meet this objective. Thus the standard-length module should aim to collect data on all the income sources covered by this chapter, whereas the shorter module need focus only on sources that are of particular interest to policymakers. (The standard module can still be used in a short questionnaire if survey planners believe that the additional information that it collects is of sufficient interest in its own right, but the short module should not be used in a standard LSMS questionnaire.)

When designers wish to measure total household income, it is necessary to distinguish between private transfers that need to be repaid and those for which no repayment is expected. In practice respondents may sometimes find it difficult to make this distinction, given the ambiguity that might be associated with having “an obligation to repay.” While in some cases

there may be no economic obligation to repay, respondents may nonetheless feel a moral obligation to do so.

Nevertheless, this distinction is important in measuring total household income. If a household receives a transfer that it does not have to repay, the transfer should be regarded as income. If the household receives a transfer that must be repaid, this is effectively a loan and should be regarded not as income but as part of the household's capital account.

#### *What Information on Private and Public Transfers Income Interests Policymakers and Analysts?*

Information on public and private interhousehold transfers is of considerable interest to policymakers and analysts, whether or not it is used to measure total household income. But how detailed does transfer income information need to be? The answer to this question depends on how survey data are likely to be analyzed. If the incidence and impact of government transfers will be studied, information on transfers should be collected at as disaggregated a level as possible, with clear distinctions made among the different types of transfers (such as child allowances and unemployment benefits). As noted above, collecting such information at a more disaggregated level also has the benefit of reducing the likelihood that a household's overall receipt of government transfers will be underestimated because respondents neglect to mention items about which they are not explicitly asked.

The nature of private interhousehold transfers and the reasons for these transfers are usually harder to identify than is the case for government transfers. Thus for a private interhousehold transfer it is desirable to collect information not only on the amount transferred but also on the characteristics of the donor, the reason for the transfer, and whether the transfer is a one-off or a regular occurrence. The benefits of this information include enabling analysts to assess whether or not these private transfers are a reliable source of income for the household.

The questionnaire should collect data on private transfers both to and from households. (Chapter 5 on consumption collects information on household expenditure on transfers.)

#### *Should the Questionnaire Collect Information on a Household's Family and Social Networks?*

Collecting information on the social networks to which households belong helps analysts model the

determinants of private transfer flows and understand why some households receive or make such transfers while others do not.

Collecting information on social networks may have extensive implications for the design of the questionnaire. Some previous surveys have collected such information—notably the Cartagena survey studied by Bamberger and others (1991). In the Cartagena survey this involved collecting information on (nonresident) close relatives, distant relatives, and friends of the household head and the head's spouse or partner. Questions were asked not only about transfer flows between the household and these people, but also about the characteristics of these people (whether or not they gave or received transfers). There were questions about where these people lived, how their standard of living compared with that of the sample household, and the frequency with which members of the sample household visited them. The questions yielded information on the size of a household's social network, the likelihood of the household receiving transfers from members of its social network (or making transfers to these members), and the extent of the household's efforts to maintain its network (through visiting).

Having this kind of information helps analysts understand why some households give or receive private interhousehold transfers while others do not. However, even when this information is available, it can be difficult for analysts to get a clear and complete picture of the extent of a household's network (especially of household friends) and to discover to what extent and in what circumstances a household can rely on receiving transfer income from members of the network.

It is also possible to use a simpler questionnaire design to gather some information on the likelihood of a household receiving or making private transfers. The draft household roster module introduced in Chapter 6 (and presented in Volume 3) collects information on the parents and children of household members who live in different households. If this information can be matched with the information collected in the transfers income module on the donors of such transfers, it can be used to analyze the determinants of private interhousehold transfers. This information will give analysts an idea of the extent to which close family provides support to poor or vulnerable households. However, an analysis based on this information alone cannot fully explain why some households receive (or make) pri-

vate transfers while others do not; nonresident parents and children form only part of a household's social network. Moreover, the household roster is not the natural place to ask questions about how close members of the sample household are to their nonresident parents or children. If such information is collected, it is almost certainly better collected in the transfers and other nonlabor income module than in the roster module. A final difficulty is that when close family members outside the household do not provide support to poor or vulnerable households, it is difficult to determine the reasons why.

***Should the Data be Collected at the Level of the Household, the Individual, or Both?***

Whether to collect some or all income data at the individual level has already been discussed in the first section of this chapter and is discussed in Chapter 17 on measuring total income. There are several potential analytical benefits to collecting data on an individual basis. The issue of who receives—and thus in principle controls—government transfers may be particularly interesting to policymakers. Also, collecting such data at the individual level may make it possible to model the responses of individuals or households to receiving this transfer (although the difficulties in accurately modeling these responses should not be understated).

When survey designers decide to collect income data at the level of individual household members, this should—to the extent possible—be done for all of the income sources covered by this chapter. In principle it may be more feasible to identify some of these income sources with individual recipients than it would be to do so with, for example, self-employment income. When the questionnaire is designed to collect the data necessary to measure total household income, it should aim to collect information on whichever household member is the recipient of each category of household income. However, survey designers should carefully consider the logistics of survey design to make sure that interview time is not wasted asking respondents irrelevant questions (such as asking young working household members about income from pensions). This practical consideration may limit the extent to which all miscellaneous income can be attributed to individual household members, but even where such an exact attribution cannot be made it is usually still possible to find out which household members benefited from a given income source.

***Should this Module Collect Information on Revenue from the Sale of Assets?***

It was suggested in the first section of this chapter that the transfers and nonlabor income module may be a good place to collect information on the household's revenue from the sale of its assets to the extent that this information is not collected elsewhere in the questionnaire. For most purposes, all that needs to be collected is information on revenue from the sale of each category of assets; collecting such information is likely to be possible in both the standard and shortened versions of the proposed module. However, as noted above, this is one of several areas where close coordination is needed between the miscellaneous income module and the other modules in the questionnaire to ensure that information on revenue from the sale of each category of assets is collected only once in the questionnaire.

**Draft Module**

The draft modules for transfers and other nonlabor income, presented in Volume 3, are based on the discussion in the first two sections of this chapter. Both standard and short versions of the module are presented. A distinction is made in the modules between private interhousehold transfers and the remaining miscellaneous income components, which are collectively referred to as “other nonlabor income.” This distinction follows common practice in many previous LSMS and other surveys (Box 11.2) and it also reflects the fact that more and different information is likely to be required (and can be collected) on private interhousehold transfers than on any of the other income sources.

***Submodule on Income from Private Interhousehold Transfers***

An important issue in collecting data on income from private interhousehold transfers is whether to collect such data from each individual household member or from a single well-informed member such as the household head. The argument can be made either way. Simply asking the household head can carry the risk that certain transfer receipts may not be mentioned because they were not received by the household head or because he or she is not aware of their existence. However, problems can also arise if each household member is questioned individually about

**Box 11.2 Cautionary Advice**

- *How much of the draft module is new and unproven?* The short version of the proposed module is largely the same as the module used for miscellaneous incomes in past LSMS surveys. In the modules presented here more detail is added in the standard version in terms of the number of income sources identified. The proposed standard version also gets a good deal more information at the individual level rather than the household level, and asks more questions about donors of private interhousehold transfers.
- *How well has the module worked in the past?* Past versions of the module have probably underestimated flows. In the past, items such as rental income, interest income, and dividends were asked about in little detail; to make matters worse, these items are not very important to the poor and notoriously difficult to collect from the wealthy. Several past surveys had so little detail about public and private transfers that some flows were probably missed.
- *Which parts of the module most need to be customized?* The categories of income to be included and the terms used for these categories must be reviewed carefully for each country. To write effective questions about government transfer programs, it is necessary to know a good deal about how payments are made (periodicity, targets and recipients, whether they vary by case or use a flat rate) for each program.

receipt of private transfers. While some transfers are made to specific members of the household, others may be made to the household as a whole, and household members may forget to mention these transfers when questioned individually. Moreover, questioning household members individually is almost always more time-consuming than putting a single set of questions to one household respondent. In most cases the costs of questioning all household members more than outweigh any possible benefits from such questioning. Therefore, the modules presented in this section are—like several other draft modules in this book—designed to be administered to one well informed respondent per household.

In both the standard and short versions of the module, to ensure the accuracy of measurements of a household's income from private interhousehold transfers, it is best to ask questions about each transfer in turn. This has the added advantage that data on the nature and characteristics of each transfer are likely to be analytically useful in their own right. Thus, follow-

ing the practice adopted in most previous LSMS and other surveys, the draft module collects information on each separate source of private interhousehold transfers (rather than asking about transfers in the aggregate).

Separate questions are included in the draft module on transfers received in cash and transfers received in kind. While this distinction may be of some interest to analysts, the primary motivation for making the distinction is that it is likely to prompt the respondent to mention all in-kind transfers received by the household. However, *great care must be taken not to include in-kind information twice in the estimate of total household income*, as the consumption module also collects information on the household's receipt of in-kind gifts of food and non-food commodities (see Chapter 5 on consumption). Total household income should generally be calculated using the data on in-kind transfers in the consumption module rather than the data in the miscellaneous income module, because the consumption data are likely to be more comprehensive.

So why include a question on in-kind transfers in the private interhousehold transfers submodule? Its purpose is to facilitate analysis of the private transfers themselves. Asking respondents about in-kind (as well as cash) transfers in this submodule means that information on the characteristics of these transfers can be collected that is not available from the consumption module. Second, it was noted above that analysts and policymakers are predominantly interested in net flows of private interhousehold transfers. Thus it is necessary to collect exact parallel information on the monetary amounts of the household's expenditure on and receipts of private transfers, and the consumption module is the only place in the LSMS household questionnaire where information is collected on the in-kind transfers made by households.<sup>8</sup>

It is desirable to collect information on the characteristics of people who give private interhousehold transfers to the respondent household. The respondent is initially asked to specify the names of the different donors simply for the purpose of identifying each transfer; this information is of no analytical interest and need not be recorded. The respondent is then asked a series of additional questions about the characteristics of the donors, such as their relationship (if any) to the household, their gender or genders, and their places of birth. Where the donors are nonresident parents or children of household members, the

responses are collected in such a way that they can be matched with the information collected on these people in the household roster module.

Information on the nature of transfers received and the motivations of people providing the transfers can be of considerable interest to policymakers. In the draft standard submodule the respondent in the recipient household is asked whether the transfer was made for a specific reason and, if so, what this reason was. The respondent is also asked whether the donor makes transfers to the household regularly or only infrequently. These questions provide some basic information on the motivation behind the transfers and allow analysts to assess how reliable a source of income the transfers are for the household.

Because the standard submodule is designed to enable analysts to estimate a household's total income, it is necessary to make a clear distinction between current transfers from other households (which the household has no obligation to repay) and capital transfers or loans made from other households (which the household is expected to repay). There are two ways in which this distinction could be made. One is for the interviewer to explain to the respondent before administering the submodule that the questions only concern transfers that the household is not obliged to repay and that any transfers that the household does have to repay should not be reported in this module. However, the problem with this is that the notion of an "obligation to repay" is not clearly defined; as noted above, respondents may feel a moral obligation to repay the transfer even if they are under no economic obligation to do so. Asking respondents to confine their replies to those transfers that they do not have to repay may lead them not to mention some transfers that should be included as a component of household income.

An alternative is for the interviewer to ask the respondent whether there is any obligation to repay the donor of a transfer directly after the respondent has reported the transfer. This procedure enables analysts to distinguish between current and capital transfers.<sup>9</sup> Thus this is the method used in the draft standard submodule (following the practice used in some previous LSMS surveys).

For each transfer, the submodule is designed so that the interviewer asks whether it was given to (or intended for) a specific individual and, if so, the identity of the principal recipient. These questions aim to

find out how private transfers affect the intrahousehold distribution of income without asking each individual household member about his or her receipt of private transfers. The submodule allows for the possibility that some transfers may be made to the whole household and not to a specific individual.

The standard submodule does not attempt to collect information on the social networks to which households belong and thus on the sources from which they may receive private interhousehold transfers. The only "social" information collected is who gave the household any transfers that they received during the reference period (in this case, the previous 12 months). As noted above, the household roster module (introduced in Chapter 6) collects information on the nonresident parents and children of household members. This can be matched with the information from the standard private interhousehold transfer submodule to establish the extent to which these nonresident parents and children of household members give transfers to the household in question. However, this yields only limited information that can be used to analyze why these nonresident parents and children did or did not provide transfers to the household. Moreover, the information is available only on nonresident parents and children of household members and not on other potential members of the household's social network, such as more distant relatives or friends. This further limits the extent to which determinants of private transfers between households can be modeled.

To collect information on a household's family and social networks, the submodule would have to be substantially different and much longer. Designing such a module would be easier for a survey conducted in a specific locality than for a nationwide survey because the nature of these social networks probably varies—for example, between urban and rural areas—as does the type of information that would need to be collected. Nevertheless, in certain instances, especially where surveys are conducted in specific localities, survey planners may wish to collect information on a household's family and social networks. This would allow analysts to study not only the influence of social networks on private transfers but also the extent to which households use such networks to gather information for finding employment or for establishing or developing household enterprises. Such inquiries are beyond the scope of the draft standard submodule pre-

sented here, and are probably best designed on a case-by-case basis.

The draft submodule on private interhousehold transfers, while slightly longer than similar modules used in some previous LSMS questionnaires, is still not very long in terms of the time needed to administer it. The additional questions in this submodule that are not used in previous LSMS and other surveys are intended to yield more information on the nature of private transfers (such as motivations behind them, frequency with which they are given, and intended beneficiaries) and to enable analysts to match the data on nonresident parents and children of household members that are gathered in this module with the data on these people gathered in the household roster module. The inclusion of these questions is justified by the fact that they usefully enhance analysts' understanding of interhousehold private transfers.

The short version of this submodule has the same basic structure as the standard version but omits several less important questions. Like the standard submodule, the short submodule collects data on household income from private interhousehold transfers, to be analyzed in conjunction with corresponding data from the consumption module on household expenditure on private interhousehold transfers. However, the short submodule does not ask respondents whether a given transfer must be repaid (a difficult question for households to answer in any case), since any survey that included the shorter version of the submodule would not be aiming to collect the data necessary to measure total household income. The short submodule gathers little information on the nature of the transfers or their donors. Rather it focuses on measuring the magnitude of the transfers.

Finally, as has been argued above, whether the standard or shorter submodule on income from private interhousehold transfers is chosen, the survey should include an equivalent submodule on households' expenditure on such transfers. A submodule of this kind has been included with the consumption module in this volume, designed alongside and symmetrically with that for incomes from these transfers. In many questionnaires it will be most natural to have the questions on transfers received next to the questions on transfers given; this can be done by merging these two submodules into a larger module called private interhousehold transfers.

#### *Submodule on Other Nonlabor Household Income*

The submodule on other nonlabor household income covers all of the components of household income not discussed elsewhere in the questionnaire. A number of issues need to be considered in adapting this submodule to a particular survey situation. One of the most fundamental issues is whether information on these other sources of household income should be collected at the individual level, at the household level, or in some combination of the two. The other nonlabor income submodule collects some information at the individual level, although some incomes are only estimated at the household level. The submodule could easily be modified to collect more income information at the individual level, but implementing this in practice might not be straightforward. The short submodule collects all information at the household level and does not attempt to attribute any of the receipts to individuals. As argued above, it is unrealistic—and of limited analytical interest—to attempt to collect all information on transfers and nonlabor income at the individual level.

A second issue is whether data on household's receipts from the sale of capital assets should be collected in this submodule. Both the standard and shorter versions allow for this. As argued above, the extent to which these categories of capital receipts are included in this submodule will depend in each case on the extent to which this information is collected elsewhere in the questionnaire. The draft submodules introduced here aim simply to list the categories that should be included somewhere in the questionnaire. The designers of the questionnaire need to be careful to avoid omissions as well as double-counting.

A third issue is how many categories of other nonlabor income should be covered. The standard submodule presented in this book covers all such categories because it is part of a questionnaire designed to collect the data necessary for measuring total household income. However, the short submodule does not need to be so comprehensive.

Related to the third issue is another issue: what the categories of other nonlabor income should be, irrespective of whether the standard or shorter submodule is used. This is very difficult to specify across the wide range of countries and contexts in which an LSMS survey might be conducted, for two reasons. First, the categories and amounts of other nonlabor income that households receive vary considerably

from one country to another. In many transition countries, government transfers of various kinds can account for a large proportion of household income,<sup>10</sup> but these transfers are a relatively insignificant source of household income in many of the poorer African countries. Such differences have obvious implications for the level of detail in which questionnaires should collect information on these categories. Second, the form taken by any one category of other nonlabor income may vary from one country to another. Government transfers might be given in many different ways including child benefits, food stamps, or school feeding programs. Thus survey designers must take care to customize questionnaire to reflect prevailing local circumstances.

Given the country-specific nature of many of these issues, this chapter does not seek to outline a precise design for the other nonlabor income submodule, nor does it aim to provide a definitive list of income categories that the submodule should cover. Instead, the chapter aims to provide guidance so that survey planners can choose the categories and the overall design that are most relevant to the country of the survey.

The standard version of the other nonlabor income submodule consists of two parts, which can—but do not have to—follow one another in the questionnaire.<sup>11</sup> The first (Part B1) collects information on relatively frequent and regular transactions that can fairly easily be attributed to individuals. While these are mostly government transfers, private pensions and other categories of frequent transactions may also be included. Information is collected on which individuals within a household receive a given category of income and on how much income they obtained.

Part B2 of the standard submodule collects information from one selected respondent per household on irregular or infrequent income sources. Income data are collected at the household level—the only practical option for these types of income—but a question is included asking whether a given receipt of income can be attributed to any specific household member and, if so, which member was the principal recipient or intended beneficiary.

Which transfers should be addressed in Part B1 of the submodule and which should be addressed in Part B2 will vary from case to case. Part B1 should collect information on transactions that are both more regular and more easily attributed to individual

members—which are also the transactions on which analysts will most likely need individual-level data. Where government transfers and other transfers are very significant income sources, it may be desirable to extend the first part of this submodule. Where such income sources are insignificant, it may be appropriate to omit Part B1 of the submodule and simply include these sources as income categories in Part B2.

Box 11.3 lists a range of possible income categories that could be included in this submodule and suggests in which part they might most naturally be included. Customizing the submodule for the country studied may require dropping some categories and adding others. In many instances customization requires renaming categories to make them meaningful in the local context (for example, using the precise local name for categories of government transfers).

The short version of the other nonlabor income submodule covers a limited number of income categories, and covers them only at the household level. Like the short version of the submodule on private interhousehold transfers, it omits the question about which household member was the primary recipient of a given transfer (as respondents may find it difficult to answer this question or may not provide reliable responses). Box 11.4 sets out the categories that might be covered in a short submodule on other nonlabor income; as before it is critically important that survey planners customize the module to suit specific circumstances of the country studied.

### Notes and Comments on Draft Modules

This section provides annotations to specific questions within the draft modules on transfers and nonlabor incomes introduced in the previous section and presented in Volume 3. The question numbering used is for the standard versions of the questionnaire.

#### *Private Interhousehold Transfers Submodule*

The vast majority of LSMS surveys have included modules that collected data on income received by households through private interhousehold transfers. The format of many of these modules has been similar to the format of the draft short submodule introduced in the previous section (and presented in Volume 3). However, some of the questions in the draft standard submodule have only rarely been used in previous LSMS and similar household surveys, including:

**Box 11.3 List of Other Nonlabor Income Categories**

The following are income categories that could be included in the standard submodule:

*Transfers*

- State pension.\*
- Private and company pensions.\*
- Survivor's pension.\*
- Unemployment benefits.\*
- Illness and disability payments.\*
- Child allowances (such as child care benefits and childbirth benefits).\*
- Job search programs.\*
- Educational scholarships (only where not collected in the education module).\*
- Maintenance payments.\*
- Dowry.
- Supplementary feeding schemes.
- Transfers from churches/mosques/other religious organizations.
- Insurance payments.
- Food or meals from NGOs.
- Inheritance.

- Official development programs.
- Zakat (a form of Islamic charity intended to benefit the poor).
- Other transfers from NGOs (not credit).

*Rental income*

- Income from renting out dwellings.
- Income from renting out land.
- Income from renting out equipment.
- Income from renting out consumer durable goods.

*Revenue from sale of assets*

- Revenue from sale of land.
- Revenue from sale of livestock.
- Revenue from sale of durable goods.
- Revenue from sale of dwellings and buildings.

*Other*

- Interest on savings.
- Dividends.
- Lottery winnings.
- Other gambling income.
- Other income (specify).

\* Categories on which data can be collected in Part B1 of the standard submodule. (Information on the other categories should be collected in Part B2 of the standard submodule.)

Source: Author's summary.

**Box 11.4 Categories of Other Nonlabor Income to Be Included in the Short Submodule**

The following are categories that could be covered in a short submodule:

- State pensions.
- Company/private pension or retirement fund.
- Social security payments (specific to country studied).
- Employee welfare schemes.
- Medical, life, unemployment insurance.
- Maintenance payments.
- Dowry and inheritance.
- Transfers from religious organizations (country-specific).

- Additional transfers from nongovernmental organizations (including in-kind transfers).
- Income from leasing assets: land, dwelling, buildings, productive assets, durable goods.<sup>a</sup>
- Revenue from sale of land.
- Revenue from sale of productive assets.
- Revenue from sale of jewelry.
- Revenue from sale of dwellings.
- Revenue from sale of durable goods.
- Revenue from loan repayments received by household.
- Other income (specify).

a. These types of rent can be separated in some instances.

Source: Author's summary.

- Questions aiming to identify the purpose of a transfer.
- Questions on the frequency of private transfers from a given donor.
- Questions aiming to identify to which household member a transfer was sent.
- Questions that identify donors who are nonresident parents or children of a household member to

match with data on the same people in the appropriate sections of the household roster module. These questions are not completely untried; the LSMS survey conducted in Kagera, Tanzania included the first type of question, and the LSMS survey in Pakistan included the second and third types of questions.

Respondents may feel more reluctant to answer some of these questions (for example, the second and third questions above) than others. In some cases (such as the first and third questions) there may not be a clear answer; the design of the questionnaire allows for this eventuality. Many of these sensitive questions are omitted in the short submodule.

It is vital in either version of the submodule on income from private interhousehold transfers to customize the selection and wording of the following types of questions for a given country situation:

- Questions that ask where the donor lives (Q12 and Q13 in the standard module).
- Inquiries into the motivation for providing the assistance (Q18 in the standard module).

In countries where transfers are widespread and frequent, it may be desirable to use a shorter reference period. (For example, in the Kagera, Tanzania LSMS questionnaire a six-month reference period was used—in part because households were to be interviewed every six months.) However, for most households the receipt of private transfers is likely to be an infrequent or irregular event, meaning that survey designers should choose a longer reference period.

**Q2.** The list of all donors should be completed before the interviewer asks questions 3–19 about each transfer.

**Q8–Q9.** Where a transfer has been made by a household member's nonresident parent or child, these questions are designed so the data they collect can be matched with data collected in Parts B and C (respectively) of the household roster module. This enables analysts to identify individuals named in Parts B and C of the household roster module who did not send any transfers to the household during the previous 12 months. It must be clear that the relationship specified in question 6 is the relationship of the donor to the recipient, and not the other way around.

**Q12–Q13.** As noted above, the options for the responses to these questions must be country-specific.

**Q15.** The example of in-kind assistance used in this question—in this case, food—is likely to vary in different countries depending on what forms of in-kind assistance are most common.

**Q18.** The options for responses to this question must be country-specific.

At the conclusion of this submodule respondents should be asked if they received any other private transfers from other households that they have not already reported. If so, the names of the donors should be listed under question 2, and the interviewer should pose questions 3–20 to respondents about these additional transfers.

#### *Other Nonlabor Income Submodule*

Almost all previous LSMS surveys have included an other income submodule. In most cases these modules have had a format very similar to the format in the draft short submodule introduced in this chapter (and presented in Volume 3). The types of income sources covered in these submodules have also been very similar to those suggested in the short module, including capital receipts. However, the format of the draft standard submodule presented in this chapter differs from that used in many past LSMS surveys, as it aims to collect more information at an individual level by identifying individual beneficiaries where possible—especially for the transactions covered in Part B1. The format introduced in the standard submodule is not totally new, however; some past LSMS surveys, including ones conducted in South Africa, Tanzania (Kagera), and Ecuador, have questioned individual household members about some of these other income sources.

Two issues are very important in the design of the other nonlabor income submodule. First, when the standard version is used, it is important for survey planners to evaluate whether government transfers and other frequent transfers are of sufficient importance in the country of the survey to justify including Part B1 of this submodule. Where such income sources are relatively insignificant, Part B1 can be omitted, and information on income from these sources can be collected by adding further categories to Part B2. Where the income sources covered by Part B1 of this submodule are highly significant, survey planners may wish to extend the draft submodule.

Whether the standard or shorter submodule is used, it is critically important that survey planners choose the categories of other nonlabor income most appropriate for the country surveyed. These are likely to vary significantly from one country to another, particularly as sources of income are likely to have differ-

ent names in different countries (for example, different types of government transfers). The questionnaire in the standard submodule should attempt to cover as many of these potential income sources as possible. Thus the lists provided in Boxes 11.3 and 11.4 should be thought of as no more than starting points.

#### **REGULAR AND FREQUENT INCOME SOURCES (PART B1).**

The list of income sources from which respondents are asked to report their income should include the items marked with an asterisk in Box 11.3 that are applicable to the prevailing circumstances in the country of the survey. The list of income sources should also include any similar income that may be received regularly or frequently and that is paid predominantly to individual household members.

In administering this submodule, it is important to be sure that payments for or to children are not omitted. Questions about child benefits paid to parents should be included here. Where benefits are provided directly to children (such as milk rations given to young children) respondents should be asked a similar set of questions about these benefits. As the questions (notably question 3) would need to be slightly differently worded, an additional submodule might be needed to cover benefits, with a design very similar to that of Part B1 of the standard other nonlabor income submodule.

**LESS FREQUENT INCOME SOURCES (PART B2).** The income sources covered by this submodule have not been listed explicitly in the draft module but should include items not marked with an asterisk in Box 11.3 that also apply to the prevailing circumstances in the country of the survey.

Where this submodule is included in a short questionnaire (which does not aim to collect the data necessary to estimate total household income), the sources covered here should include items in Box 11.4 that apply to the prevailing circumstances in the country of the survey. Similar issues arise in designing the shorter other nonlabor income module. Here the most important issue will be to ensure that all the most important sources for the country in question are covered.

#### **Notes**

The author is grateful for the very helpful comments of Donald Cox, Paul Glewwe, Emmanuel Jimenez, Margaret Grosh, and an anonymous reviewer on an earlier draft.

1. As discussed above, imputed rent from owner-occupied dwellings is not included here. This source of rental income is likely to accrue to poor households as well as to rich households.

2. Identifying a vulnerable household is both more difficult and more subjective than identifying a poor household, especially based on a single cross-section of data. This issue is discussed by Glewwe and Hall (1998) using panel data for Peru.

3. There are many possible reasons why a considerable number of poor households do not receive private interhousehold transfers. Poor households may not have relatives elsewhere, their relatives may be as poor as they are, or there may have been a breakup of the family.

4. This qualification is important. In some countries, large volumes of private interhousehold transfers may cross international boundaries—for example, in the case of migrant workers from southern African countries working in South Africa or migrant workers from some Asian countries working in the Middle East. In such circumstances, the comparison suggested in this paragraph cannot be made.

5. This may not always be the case. If the donor feels the recipient is spending the transfer in a wasteful or inappropriate way (for example, on increased consumption of alcohol), the donor may choose to withdraw or reduce the transfer.

6. Some information on asset sales is likely to be collected elsewhere in the questionnaire. Sales of agricultural equipment are recorded in the standard version of the agriculture module and sales of land in the expanded version. Sales of household business assets are recorded in the standard version of the household enterprise module. However, given that sales of assets may not be covered comprehensively elsewhere in the questionnaire, the miscellaneous income module is probably the most appropriate place to ask questions about the sale of assets (such as consumer durables) that were not covered elsewhere.

7. It may not be easy to distinguish households that sold their assets to finance consumption from households that sold one type of asset to acquire another type. However, it should be possible to make such a distinction when the questionnaire also contains corresponding information on asset purchases (available in the standard modules for most categories of assets).

8. It may not be as important to collect such detailed information on the nature and characteristics of outgoing transfers as on incoming ones, since incoming transfers are generally of greater interest to analysts. While the questionnaire in Volume 3 presents similar modules on incoming and outgoing interhousehold transfers, the section on outgoing transfers may be shortened if need be.

9. Experience from previous LSMS surveys suggests that where the distinction has been made, the vast majority of private interhousehold transfers reported by respondents have been current rather than capital transfers.

10. In Bulgaria the 1995 Integrated Household Survey found that the average proportion of household income coming from state transfers was 37.3 percent. More than 10 percent of the sample derived all of their measured income from this source.

11. This design combines elements of designs used in some previous LSMS questionnaires (including questionnaires used in Kagera, Tanzania and in Pakistan) with some new elements.

## References

- Bamberger, Michael, Daniel Kaufman, Eduardo Velez, and Scott Parris. 1991. "Intrahousehold Transfers and Survival Strategies of Low-Income Households: Experiences from Latin America, Asia, and Africa." World Bank, Washington, D.C.
- Corbett, Jane. 1988. "Famine and Household Coping Strategies." *World Development* 16 (9): 1099–112.
- Cornia, Giovanni Andrea, and Frances Stewart. 1995. "Food Subsidies: Two Errors of Targeting." In Frances Stewart, ed., *Adjustment and Poverty: Options and Choices*. London: Routledge.
- Cox, Donald, and Emmanuel Jimenez. 1990. "Social Objectives through Private Transfers: A Review." *World Bank Research Observer* 5 (2): 205–18.
- . 1992. "Social Security and Private Transfers in Developing Countries: The Case of Peru." *World Bank Economic Review* 6 (1): 155–69.
- . 1995. "Private Transfers and the Effectiveness of Public Income Redistribution in the Philippines." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence*. Baltimore, Md.: Johns Hopkins University Press.
- . 1998. "Risk Sharing and Private Transfers: What about Urban Households?" *Economic Development and Cultural Change* 46 (3): 621–37.
- Glewwe, Paul, and G. Hall. 1998. "Are Some Groups More Vulnerable to Macroeconomic Shocks than Others? Hypothesis Tests based on Panel Data for Peru." *Journal of Development Economics* 56 (1): 181–206.
- Greene, William H. 1993. *Econometric Analysis*. Second ed. Englewood Cliffs, N.J.: Prentice Hall.
- Grosh, Margaret E. 1992. "The Jamaican Food Stamps Program: A Case Study in Targeting." *Food Policy* 17 (1): 23–40.
- . 1995a. "Towards Quantifying the Trade-off: Administrative Costs and Incidence in Targeted Programs in Latin America." In D. van de Walle and K. Nead, eds., *Public Spending and the Poor: Theory and Evidence*. Baltimore, Md.: Johns Hopkins University Press.
- . 1995b. "Five Criteria for Choosing Among Poverty Programs." In N. Lustig, ed., *Coping with Austerity: Poverty and Inequality in Latin America*. Washington, D.C.: Brookings Institution.
- Haddad, Lawrence, and Manfred Zeller. 1997. "Can Social Security Programs do More with Less? General Issues with Challenges for Southern Africa." In Lawrence Haddad, ed., *Achieving Food Security in Southern Africa: New Challenges, New Opportunities*. Washington, D.C.: International Food Policy Research Institute.
- Heckman, James J. 1979. "Selection Bias as Specification Error." *Econometrica* 47 (1): 153–61.
- . 1990. "Varieties of Selection Bias." *American Economic Review, Papers and Proceedings* 80 (2): 313–18.
- Jarvis, Sarah J., and John Micklewright. 1995. "The Targeting of Family Allowances in Hungary." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence*. Baltimore, Md.: Johns Hopkins University Press.
- Kanbur, Ravi, Michael Keen, and Matti Tuomola. 1995. "Labor Supply and Targeting in Poverty-Alleviation Programs." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence*. Baltimore, Md.: Johns Hopkins University Press.
- Knowles, James C., and Richard Anker. 1981. "An Analysis of Income Transfers in a Developing Country: The Case of Kenya." *Journal of Development Economics* 8 (2): 205–26.
- Quisumbing, Agnes R., Lynn R. Brown, Hilary S. Feldstein, Lawrence Haddad, and Christine Peña. 1995. *Women: The Key to Food Security*. Washington, D.C.: International Food Policy Research Institute.
- Sahn, David E., and Harold Alderman. 1995. "Incentive Effects on Labor Supply of Sri Lanka's Rice Subsidy." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence*. Baltimore, Md.: Johns Hopkins University Press.



# 12 Housing

*Stephen Malpezzi*

Housing characteristics, and the process by which housing is constructed and occupied, are key aspects of the living standards of households in developing countries. Housing is of great importance to households in both developed and developing economies, because it is the largest fixed capital investment that households make. In developing countries, housing accounts for 10–30 percent of household expenditure, 6–20 percent of GNP, and 10–50 percent of gross fixed capital formation. Furthermore, as economies develop, the proportion of GDP accounted for by housing investment rises. Other than human capital, housing and land are the types of capital that are most widely owned.

There are three main ways that housing data are used in policy research and thus three reasons why housing data should be collected in LSMS surveys. First, housing information provides useful direct indicators of living standards, including access to electricity and clean drinking water, type of dwelling, toilet facilities, and living space per person. Second, housing is a form of consumption that can be overlooked when analysts estimate overall standards of living using household survey data. For example, families that rent their housing report their rent payments as part of their overall expenditures, whereas families that own their housing often report incurring little current expenditure on housing—as they are consuming the fruits of a previous investment. Thus estimates of total household consumption should include the implicit rent of owner-occupied housing. Third, housing data can be used to understand why particular housing conditions exist and whether specific government policies can be adopted that will lead to more efficient or more equi-

table outcomes. As is explained further in the first section of this chapter, governments regulate and intervene in housing markets in many ways, and household survey data can be used in analyses that determine the effectiveness of these policies.

This chapter discusses what policymakers need to know about housing and housing markets and which housing issues can be analyzed using data from household surveys such as the Living Standards Measurement Study surveys. The first section of this chapter discusses key housing policy issues and shows how housing market analysts can address these issues. The second section reviews the data that would need to be collected in a multitopic household survey to make it possible for these issues to be analyzed. The third section contains a draft prototype housing module that can be customized to match the prevailing conditions in the country of the survey. The fourth section provides explanatory comments on the draft module.

## Housing Policy Issues

This section discusses in detail the ways in which data on housing collected in multitopic household surveys like the LSMS surveys can be used to analyze some key issues in the housing sector. Box 12.1 reflects this discussion in that it shows which issues can and which cannot be analyzed with LSMS-type data.

### *Using Housing Characteristics as Indicators of Living Standards*

In order to use the characteristics of a household's dwelling as indicators of the household's standard of living, analysts require data on those housing characteristics. Exactly which of these characteristics are useful is discussed in the next section of this chapter.

### *Measuring Housing Consumption*

A second reason for collecting housing data is to obtain the information needed to derive a correct estimate of a household's consumption of housing. In principle, households purchase accommodation (or produce it for themselves) just as they purchase food, clothing, and other consumption items. As explained in Chapter 5, total consumption is a crucial indicator of household welfare, so it is important that it be cal-

culated carefully. For most common purchases, such as purchases of food and clothing, the cost of the items is their market price, which is the value that should be placed on these items. However, some households' housing may not be purchased, or even rented, in a directly observed transaction at its true market price. For example, some housing is inherited, and some is built by the households themselves. Some households that rent housing do so at subsidized or controlled prices. Therefore, to measure housing consumption correctly, it is necessary to use market prices or an estimate of such prices.

Another problem is that households that own their housing incurred much of the cost many years ago but still use the dwelling today. Yet if two households live in similar dwellings, their standard of living is similar regardless of when the housing was purchased. Thus it is necessary to estimate each household's current consumption of housing by estimating what the household would spend to rent an equivalent unit at market prices. Therefore, the household survey should collect data on market rent (if observed) or this should be estimated (if not directly observed) for each survey household's housing unit.

Virtually every housing unit is unique in terms of its size, quality, location, and other characteristics. This

#### **Box 12.1 Policy Issues and Housing Data**

*Policy issues that can be analyzed with cross-sectional data from LSMS-type surveys*

- The level and distribution of housing consumption
- The distribution of housing assets
- The frequency and distribution of specific housing characteristics and conditions (such as space, sanitation, age, condition, and crowding)
- Housing tenure, tenure security, and tenure choice
- The demand for housing
- The determinants of price changes
- Upgrading
- The measurement and determinants of vacancies
- The valuation of housing subsidies (for example, from public housing, or from rent-controlled markets)
- How households finance their housing
- Behavior relating to housing finance and savings
- Satisfaction with the neighborhood and the unit

*Policy issues that can be analyzed with panel data from an LSMS survey*

- Housing filtering (changing supply from the existing stock)

- The determinants of price changes (panel data better than one cross-section of data)
- Tenure choice (panel data better than one cross-section of data)
- Upgrading (panel data better than one cross-section of data)
- Vacancies (panel data better than one cross-section of data)

*Policy issues that cannot easily or directly be analyzed with data from an LSMS survey*

*(Note: Many analyses of these issues make indirect use of some household survey data.)*

- The regulation of development (for example, zoning and building codes)
- The determinants of the supply of new construction
- Changes in the supply of serviced land
- Housing investment and the business cycle
- Net effects of government interventions on producer and consumer incentives

heterogeneity, the durability of most housing, and the many forms of tenure and payment that exist can make it a complex process to estimate market prices. Some of the relevant measurement issues are briefly discussed in Appendix 12.1; see also Green and Malpezzi 1998.

### *Understanding Housing Market Behavior*

The third reason for collecting housing data in LSMS surveys is to help analysts and policymakers understand how housing markets work and how government policies affect housing outcomes. In this chapter the discussion of housing market analysis focuses on the analysis needed for public policy purposes because of the overall purpose of this book. Nevertheless, analysis of market behavior is of interest to, for example, housing providers and academics as well as to policymakers.

In principle, government interventions in housing markets can correct for market failures and produce positive externalities for society as a whole. In most countries governments define and enforce property rights, which are the “rules of the game” and the essential element of a successful housing market. However, there is no guarantee that all public interventions will have positive outcomes in practice. There are many examples of public interventions that have exacerbated market failures as well as examples of interventions that have been more successful (World Bank 1993a). Much depends upon the capacity of the institutions in the country of the intervention and the prevailing process of housing development and management.

The remainder of this section will discuss several types of housing market analysis. Before doing so, it is useful to review the different ways in which government housing policies can affect housing outcomes, since this is the most direct way governments can improve housing conditions in developing (and developed) countries. Of course, the other obvious way that government policies can improve housing outcomes is by increasing economic growth, which raises household income, allowing households to purchase or rent better housing.

Besides the general role that governments play in providing a stable macroeconomic environment conducive to housing investment, there are many types of government policies that affect housing. The most important are:

- Assigning and enforcing property rights with respect to land and real estate, including housing.

- Regulating development by means of zoning, subdivision regulation, and building codes.
- Providing public housing, either directly or through state-owned enterprises.
- Taxing and subsidizing housing.
- Enforcing rent control and other rental regulations.
- Regulating other aspects of the real estate industry, such as construction and brokerage.
- Providing infrastructure such as electricity, water, and sewage.
- Regulating finance through interest rate regulations, the provision of credit, and the prudential regulation of lenders.

The relationship between policies and housing outcomes can be studied using both descriptive analysis and estimations of behavioral relationships. Descriptive analysis is essentially the tabular presentation of simple statistics on housing, such as which households rent or own, which live in subsidized units or units subject to rent control, how much households pay for their housing, and how dwellings were obtained (whether inherited, purchased, or built). These basic characteristics of housing can also be cross-tabulated by different income and tenure groups. This type of analysis is very useful for getting an initial snapshot of various government policies and of the general state of the housing market, but it cannot usually provide quantitative estimates of the effects of government policies on the housing market. To find out how different housing policies affect housing outcomes, analysts need to understand household behavior. The following two subsections examine these two categories of analysis—descriptive analysis and the estimation of behavioral relationships.

### *Descriptive Analysis*

Good descriptive analysis can provide policymakers with key facts about the housing market. For example, it can show which income groups benefit from subsidized housing and which households constructed their own dwellings (and thus would not be directly affected by changes in construction industry regulations). Three basic types of data are most useful for descriptive analysis: data on the housing stock, data on housing expenses (including taxes and subsidies), and data on property rights (including rental arrangements).

**THE HOUSING STOCK.** Perhaps the most obvious data to collect on housing is information on the physical

characteristics of the dwelling, some of which are basic indicators of a household's living standard. This general use of data was already considered above and will be discussed in more detail in the second section of this chapter. However, there are certain dwelling characteristics that have particular significance for policy-making. Descriptive analysis of housing stock data can be used to examine:

- The characteristics of a dwelling that yield information about the incidence of taxes, subsidies, or regulations. For example, information on the relative importance of indigenous versus "modern" construction techniques or single-family homes versus multi-family housing often provides estimates of how much housing is subject to particular regulations or taxes. It is often useful for policy-makers to know how such characteristics vary by region and income.
- Dwelling characteristics related to basic quality standards and building code requirements such as regulations concerning water supply and sanitation.
- Vacancy patterns and how these vary by location.
- In many countries, the differences in the quality of the housing between "formal" and "informal" submarkets. How do crowding, vacancies, and other market outcomes differ in these submarkets?

Malpezzi (1984, Appendix F) provides a convenient list of descriptive tables and cross-tabulations on housing stock and related variables, which can be a useful starting point for a descriptive analysis plan. Mayo and others (1982) provides an excellent illustration of how household survey data can be used to describe and analyze basic housing market outcomes such as quality and the policy implications that can be drawn from such analysis.

#### **HOUSING EXPENDITURES, TAXES, AND SUBSIDIES.**

Obviously, analysts need basic information on housing expenditures to estimate any meaningful welfare measure for households and also to analyze the issues of housing subsidies and taxation, which are discussed below. Key issues for descriptive analysis are:

- Whether households that own their dwelling are making payments on loans or mortgages, the size and term of such payments, and when the loans will be paid off.
- The amounts that renters pay and the form that rent payments take (for example, cash, in-kind, or work) and any utilities included in these payments.

- Payments by both renters and owners for utilities and other housing-related services (such as water, sewerage, electricity, and telephone services).
- The shares of the housing market that are financed formally and informally, the terms, and how these vary by income, region, and other household characteristics.
- Direct taxes paid, either by renters or owners.
- Subsidy payments received by (or payments made on behalf of) renters and owners.
- The proportion of their income that households typically spend on housing and how this varies by type of tenure, the household's income level, region, and other household characteristics.
- The consumer's surplus gains and losses from subsidies.

The notion of a consumer's surplus is important and merits a brief explanation. When a government subsidizes a household by giving it an unrestricted cash grant, a dollar is worth a dollar, a peso a peso, a ruble a ruble. However, when the government subsidizes a household by providing it with a good or service (such as housing) or if the government requires that a household spend a transfer of cash in a certain way (for example, on housing), the value of the subsidy to the household is usually less than its cash value. Measuring the household's actual benefit from such a transfer is the aim of measuring the consumer's surplus. Detailed discussions of the consumer's surplus and related concepts can be found in Green and Malpezzi (1998), Freeman (1979), and Deaton and Muellbauer (1980). On subsidies, including the application of consumer's surplus, see Kim (1991), Sanyal (1981), Mayo (1986), and Yu and Li (1985). For more general analyses of incentives that examine a wide range of such interventions, see World Bank (1989) and Malpezzi and Mayo (1997).

**PROPERTY RIGHTS.** Until the last decade, property rights in developing countries had not been analyzed in much depth, largely because they are well-established in many developed countries and have therefore been taken for granted. Nevertheless, property rights are still an issue in many developing countries, particularly in the transition economies of Eastern Europe and the former Soviet Union. The most important kinds of property rights data for descriptive analysis include:

- How many households own and how many rent, and how this ratio varies by region and income

group. For owners, information on the specific nature of property rights is also useful.

- For renters, the form of their rental arrangement, such as the length of the lease (if specified), from whom the dwelling is rented, and if there is any relationship between the tenant and the owner.
- For owners, the existence of any official title or deed for the house and for the land it is built on, exactly who owns the title to the unit (or, for renters, who signed the lease) and what kind of title it is, and the extent to which the household's ownership of the title is secure.
- The length of time the household has lived in the dwelling, whether rented or owned. If owned, how the dwelling was obtained; if rented, the details of the lease.

Additional discussion of property rights can be found in Kiamba (1989) for Africa, Bromley (1989) for Asia, and Betancur (1987) and Gilbert (1989) for Latin America. The recent literature is dominated by analysis of property rights in formerly socialist countries; see, for example, Jaffe (1993), Jaffe and Louziotis (1996), and Pejovich (1990). Examples of research on forms of housing tenure and the value of this tenure include Jimenez (1982a, 1982b, 1984) and Tipple and Willis (1991b).

#### *The Estimation of Behavioral Relationships*

The above discussion showed how simple descriptive statistics can be used to get an idea of how government policies may affect housing outcomes. However, descriptive analysis is mainly concerned with “what is.” To answer “why,” it is important to know how households (and other relevant actors like suppliers and governments) behave. For example, descriptive statistics can show how much households spend on housing on average. This description can be extended by presenting averages for, say, different income groups. However, to understand more about the underlying behavior of households and other relevant actors, analysts can go a step further and, for example, estimate the “income elasticity of demand,” a summary numerical measure of how much housing expenditure increases as income increases.<sup>1</sup> Analysis of housing markets can be complicated by many different factors, including housing's physical and locational heterogeneity, imperfect information about buyers and sellers, illiquidity, significant environmental and other externalities, and time lags in supply. Many, though not all, of these issues are discussed in this chapter. For a

more elaborate treatment see World Bank (1993a) and Green and Malpezzi (1998).

**HOUSING DEMAND.** How much people are willing to pay for housing is one of the most important characteristics of the housing market that can be examined with data from a multitopic household survey like the LSMS. As noted above, and as discussed in some detail in Appendix 12.1, housing rent (both actual or imputed) is an expenditure measure and consists of price multiplied by quantity. The majority of demand studies (including Follain, Lim, and Renaud 1980 and Malpezzi and Mayo 1987a, 1987b) examine expenditure by estimating so-called Engle relationships (for example, actual or imputed rent) or sometimes house value (the present value of rent) as a function of income, demographic variables, and so on. A smaller number of studies have decomposed housing expenditure into its quantity and price components using hedonic models (see Appendix 12.1) or models in which prices vary with the interurban location of the dwelling. Ingram (1984) and Malpezzi (1998) are examples of studies that regress some quantity measure against prices as well as other factors that influence demand, such as income and household composition. Although there are a plethora of measurement and other issues to be resolved in this research area, housing demand is generally the most thoroughly studied and best understood of the major categories of housing market behavior (Olsen 1987). Key policy issues regarding housing demand are:

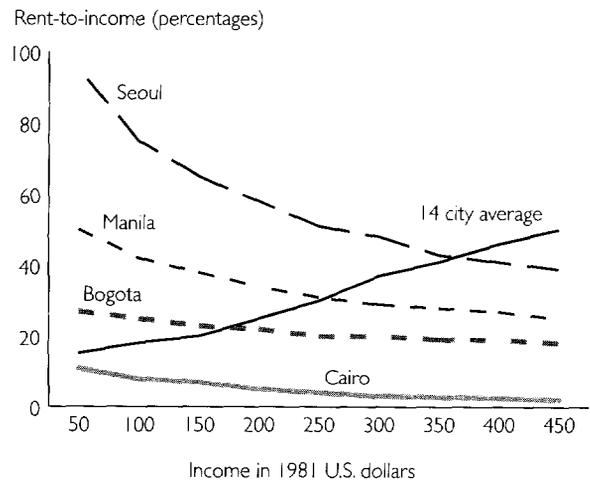
- How housing expenditures change with household income (the income elasticity of demand). Understanding this relationship is the key to understanding the often-misunderstood set of issues that are loosely labeled “affordability” issues.
- How housing expenditures change in response to changes in housing prices (the price elasticity of demand). As housing prices are affected by taxes and subsidies, this information can be used to show how tax and subsidy policies affect households' housing decisions made by households.
- How demand varies with demographic characteristics. For example, how fast does housing consumption change with household size? Do female-headed households spend more or less than average after controlling for other demand determinants?
- The determinants of demand for different tenure arrangements (owning, renting, or living in government-provided housing).

- How demand relates to the household's investment motives, as well as its demand for current consumption.
- The demand for the individual characteristics of housing such as space, quality, location, and types of amenities (such as type of toilet, drinking water, and electricity). In particular, how the location of a household's dwelling relates to the location of the workplaces of the household's members. When policymakers misunderstand the latter relationship, this can result in empty housing projects, underemployed public housing residents, and large inefficiencies in transport spending in developing, developed, and transition economies alike.

Comparative studies such as Malpezzi and Mayo (1987a) and many studies of single markets such as Follain, Lim, and Renaud (1980) have demonstrated that the parameters of demand vary from country to country in significant and at least partly predictable ways. Most studies have found income elasticities of demand that are less than 1 within markets. (In other words, housing consumption has increased with income, but less than proportionately). Despite the relative stability of within-market elasticity across countries, the average share of the household's budget that is spent on housing varies tremendously from market to market and especially across countries; see Figure 12.1 for an illustration. This relation can be examined by estimating the cross-market elasticity of average budget shares with respect to average income in each market. Malpezzi and Mayo (1987a) found that in a range of developing countries, the cross-market elasticity was actually greater than 1 (in other words, housing consumption increased somewhat faster than income).

Despite the fact that many studies have already been carried out, experience suggests it is generally worthwhile to undertake customized demand studies for a given market. There is a particular need for further research on how consumption responds to price—in other words, price elasticities (which are less settled than income elasticities). Also, much of the literature on housing demand in developing countries focuses on demand for housing as a composite good, while there is much less research on demand for individual housing characteristics such as numbers of rooms and various measures of quality. See Follain and Jimenez (1985b) for a review of the literature on demand for specific housing characteristics. Follain,

**Figure 12.1 Rent-to-Income by Income (Owners)**



Note: The authors compare markets by examining each market's median income household. The dotted lines representing Cairo, Bogota, Manila, and Seoul slope down because within markets (for example, within cities or within countries), housing consumption always increased with income but generally grew more slowly than income—in other words, the income elasticity was less than 1. Comparing 15 markets' median income (the solid line; not all 15 cities are shown), the average rent-to-income ratio in each market increases with the median income—in other words, the income elasticity is slightly greater than 1. Source: Malpezzi and Mayo 1987a.

Lim, and Renaud (1980), Ingram (1984), Mayo and others (1982), and Mohan (1994) provide useful examples of how to undertake a demand study and tailor it to specific country conditions.

**HOUSING SUPPLY.** Much less research has been done to date on housing supply, despite the fact that supply parameters are probably even more important for policymakers to know about than demand parameters. In broad terms, housing supply comes from two sources: new construction and the existing stock. Housing economists refer to changes in the existing stock as “filtering.” In common parlance, as units “filter down,” they pass from richer households (owners or tenants) to lower-income households. Units can also “filter up”—pass from poor households to richer households—if a neighborhood is being revitalized or “gentrified.” Large improvements (upward filtering) in a particular dwelling are also known as “upgrading.” (For further information on upgrading see Strassman 1982, Struyk 1982, and Rakodi 1987.) Key questions about supply include:

How much of the housing supply consists of new construction and how much is from the existing stock? How much upgrading is done in place and how

much is effective supply changed when two or more households share (or stop sharing) a dwelling?

- How does supply change in response to changes in the price of housing? What determines this elasticity? What are the effects of natural (geographical) constraints versus man-made (regulatory) constraints on supply?
- What is the role of filtering in the market? In other words, how does the supply of housing from the existing stock change to meet demand? During a given period, how much housing filters up and filters down? What are the determinants of this filtering process, and are there regulatory or other impediments to it?
- What effects do different government policies (such as rent control, the regulation of real estate industry, or government provision of housing) have on the supply of housing?
- How do these effects differ for different tenure (for example, renting versus owning), by income, and by type of housing unit?

Some supply issues are best studied with aggregate time-series data, but many can be studied with household survey data, especially if the survey has collected panel data, which would make it possible to study housing supply over time. Burns and Grebler (1977) and Renaud (1980) are examples of aggregate studies of supply. Malpezzi and Mayo (1987a) presented the first econometric estimates of supply elasticities from time-series data for several developing countries. Bramley (1993) and Ozanne and Struyk (1978) used alternative methods to study supply with household survey data. For studies of supply from the existing stock through filtering, see Green and Malpezzi (1998) for a general review and see Ferchiou (1982) and Johnson (1987) for developing country examples.

**LAND AND INFRASTRUCTURE.** Housing supply is inextricably related to the amount of land available for housing construction and to the availability of infrastructure. Major policy questions regarding land and infrastructure that require the estimation of behavioral relationships include:

- Is the supply of serviced land in urban areas expanding to meet growing population and employment needs? Which land uses are growing the fastest? Where is urban land conversion taking place? Is the supply of infrastructure keeping up with demand?

- Are land prices increasing faster than the overall rate of inflation? Where are land prices the highest and where are land prices increasing the fastest?
- How do changes in land prices affect the costs of end users? Is the price and affordability of housing and commercial and industrial space changing and are real occupancy costs greater now than before?
- Is the land market segmented—for example, divided into a formal and an informal sector? Which households do not have access to housing from the formal private sector? What regulations govern the use and sale of land?
- What is the system for providing infrastructure? What roles do the private and public sectors play in this? Are costs recovered? Does the infrastructure system respond to demand? Does infrastructure get installed in low-income areas?

Once again, many of those questions are answered most directly using aggregate or other collateral data, but many of these issues can also be analyzed using household survey data. For example, it is straightforward to add questions on land prices and land acquisition to household surveys (see Mayo and others 1982 for examples). Also, cross-tabulations of responses to household survey questions regarding services such as water, sanitation, and transport can yield insights into the provision of infrastructure. Angel and others (1986), Dowall (1991), and Farvaque and McAuslan (1992) are a few of the many useful studies of land issues. Ingram and Carroll (1981) and Mohan (1994) give particularly good accounts of the spatial structure of land markets in developing countries, and Bertaud and Renaud (1994) examine socialist countries where, until recently, land prices were not permitted to vary from place to place (or according to their productivity). Gackenheimer and Brando (1987), Lee (1992), and Lee and Anas (1992) discuss infrastructure issues in general. See Chapter 14 on the environment for a discussion of water supply and sanitation issues in great detail.

**HOUSING FINANCE.** Perhaps the most important single determinant of the quality of the housing of a given household is its income and, therefore, its ability to purchase or rent housing. Nevertheless, because all housing is an expensive and long-term investment, all housing purchases are financed in one way or another. Formal housing finance, provided by a wide variety of organizations, has been the subject of much research

in recent years. However, in many developing countries, formal housing finance institutions are relevant only to a small proportion of households. Instead, households in developing countries often turn to various informal sources of housing finance such as inter-family transfers, but these tend to be very expensive as outlined in Renaud (1984) and Malpezzi (1996). Some countries have only small enclave formal institutions that make few loans at very favorable terms. These often have little relevance to low-income and rural households whose members earn their living in the informal sector. In many respects, the challenge facing the governments of many developing countries is to encourage the development of formal sources of housing finance that are sustainable and affordable to a broad range of the population.

Given the importance of finance for determining housing outcomes, policymakers should aim to deepen housing finance markets in order to encourage investment in housing. Key issues in the area of housing finance include:

- What are the sources of housing finance, and how are these funds used? What is the system of intermediation for housing finance, and how is it connected to financial intermediation in general? What kinds of mortgage instruments are available on the lending side? What rules govern institutional features such as mortgage insurance and foreclosure?
- Are subsidies and taxes built into the financial system? If so, what is the nature and extent of these subsidies and taxes? What are the effects of tax, regulatory, and subsidy policies on the cost of credit?
- What are the mortgage interest rates, and other terms, paid by households of different types that are borrowing from formal and informal sector finance institutions? How do these terms compare to the financing available for other (nonhousing) investments, and how do they compare to inflation?
- What are the real effects of housing finance—in other words, the effects that housing finance has on housing consumption, tenure choice, and mobility? Does the availability (or lack) of formal housing finance affect such outcomes, or are formal and informal finance good substitutes?

Most research on housing finance has used institutional and macroeconomic data rather than household survey data. However, much can be learned about housing finance from household survey data if the questionnaire includes carefully chosen questions

about how the households in the sample have financed their housing and at what terms. The best example to date of housing finance analysis using household-level data is Struyk and Turner (1986).

#### *Research Methods and Data Needs*

The data needed to analyze many of the policy issues discussed above can be collected in a multitopic household survey that includes a module specifically related to housing. This housing module would gather data on, for example, housing location, housing conditions (quality and quantity), tenure, and the rents and prices that households pay. This information could then be combined with data from other parts of the household questionnaire (on, for example, household incomes and characteristics) to answer many of the questions posed in the first section of this chapter.<sup>2</sup> A well-designed housing module will also collect data that assists analysts in other ways (for example, to measure consumption accurately and precisely).

It should be mentioned at this point that there is very little information on the operation of rural housing markets in developing and transition economies. In fact, the vast majority of housing market analysis in developing, transition, and developed countries has focused on urban housing markets, thus excluding a significant slice of the housing market in the countries being studied.<sup>3</sup> While this is the case in virtually all countries, the severity of the problem that this omission presents varies from country to country. For example, in Asia, Korea is currently about four-fifths urbanized, while Thailand is about four-fifths rural. More than one-third of the populations of Poland, the Czech Republic, Hungary, Italy, and Switzerland live in rural areas. Certainly, one of the biggest contributions of LSMS surveys to housing analysis is their provision of data on rural housing markets, which can be used to research this much neglected area.

#### *Categories of Data*

This subsection outlines the categories of data that can be collected in a housing module in a multitopic household survey. It also indicates specific questions that should be included in the questionnaire, the answers to which are likely to illuminate the important policy issues outlined in the first section of this chapter.

**HOUSING CHARACTERISTICS.** The most basic data that should be collected in the housing module are data on

the characteristics of the household's dwelling. The most relevant characteristics for policy research purposes will vary somewhat from place to place, but it is always important to collect data on the basic structure of the dwelling (for example, whether it is single-family or multifamily and what material it is built with), the age of the structure, its size, the number of rooms, the number and size of bathrooms, and other characteristics related to type and reliability of its water and sanitation services.

Other important questions relate to the quality of the neighborhood in which the dwelling is located and what services are provided in that neighborhood. Not all of these location data need to be collected by asking questions of household respondents. It would be better for the interviewers to make their own observations of these phenomena while they are in a household's dwelling to conduct the survey interview. First and foremost, they should record the location of the dwelling in a city, town, or other market, since housing markets are typically analyzed by place. Within each city or market, they should indicate where the dwelling is located in relation to the central business district of a city or town. One question that must be asked of the respondent rather than observed by the interviewer is the distance household members must travel to their workplaces and the amount of time it takes them to get there. Also, it may be useful to find out how far the dwelling is from other places of employment in the area or from central locations in the metropolitan area.

**PRICES.** One set of issues that must be addressed early in the design phase of any LSMS housing questionnaire relates to measuring housing prices and consumption. These issues have been discussed briefly above and are discussed again in some detail in Appendix 12.1. Rent is the most obvious measure needed for any consumption analysis. Because rent can be observed directly for renters but not for owners, it is usually necessary to impute the rental value of an owner-occupied unit.

There are several ways to collect these data (Green and Malpezzi 1998). First, the owner can be asked how much rent they could charge for their unit. Second, the coefficients of a hedonic index estimated using a rental sample can be applied to the corresponding characteristics of individual owner-occupied units to impute rent.<sup>4</sup> A third general approach is to apply a

capitalization rate to each owner-occupied unit to appraise its value.

Appendix 12.1 briefly describes hedonic indexes and "cap rates" for readers unfamiliar with these concepts. Each approach has its pros and cons. Generally these approaches are complementary, although the hedonic approach can be especially useful. Hedonic indexes require extensive data on a unit's characteristics (such as size, type, and location) as well as on the amount of rent paid.

Whatever general approach is taken, data must be collected on arms-length market transactions, which are transactions between two parties who have no special relationship that would suggest the price paid is different from market prices. For example, transactions between close relatives may not be arms-length.<sup>5</sup> Price controls, subsidies, discounts to relatives and kin, and transactions that include in-kind rents (such as services performed in lieu of cash rent) all introduce obvious differences between the cash price paid and the arms-length market price. The questionnaire needs to differentiate households that are reporting their own rents and values based on arms-length transactions from households that are under some form of control or subsidy, are related to the landlord, and so on. A further complication is that in some markets, very few market transactions are not affected by some sort of price control. For example, in some markets, very few units are traded at market prices. This can be because housing is primarily owned by the government and is rented at very low rents (as in Moscow and, until recently, China) or because rent control is very widespread (as in Ghana; see Malpezzi, Tipple, and Willis 1991). Nevertheless, despite the problems that can be involved in interpreting such numbers in countries like Russia, it is necessary in these countries to collect data on the official (nonmarket) rent for the purposes of policy analysis.

It follows that the questionnaire should be designed to elicit from the respondent whether the household receives any housing subsidy and, if so, what kind and if the unit is subject to rent control. It is also important that the questionnaire carefully distinguish between housing and agricultural real estate in rural areas and between housing and shops, offices, and other nonresidential uses in both rural and urban areas. In addition, it should be noted whether any commercial premises are physically attached to the household's dwelling.

When the survey is fielded in countries or regions with no active housing market, it may be appropriate to include questions about housing prices in the community module of the household survey. These questions can be put to community leaders or others who are knowledgeable about what housing units exist of various standard types. These questions in the community module will supplement the housing questions asked in the household questionnaire. If the market is extremely moribund with few similar dwellings being sold, the questions included in the community module can be about the costs of constructing typical housing units.

In many countries, property taxes are an important source of government revenue (Dillinger 1991). Of course, how great a burden they impose depends on whether they are levied or enforced. In some markets, various transaction taxes and registration fees on housing sales are high. Where this is relevant, questions about such taxes and fees can easily be added to the housing module. In some markets, questions about condominium fees or maintenance fees will also be relevant.

**EXPENDITURES.** An issue that can arise when survey designers are framing the questions about housing demand is the distinction between gross and net household expenditures on housing. Some renters pay for their utilities separately from their rent, but others pay a monthly rent that includes utility charges. If more than one household lives in a unit, it is necessary for analysts to know how much money is passed from one household to another and how much goes to third parties such as the landlord. Renters may also face additional charges—particularly in controlled markets—including key money, advance rent, and expenditures on maintenance and repairs. Malpezzi (1998) discusses the role of such side payments in some detail. The questions in the housing module should cover all of these possible extra charges.

**MOBILITY.** Research has demonstrated that the longer a household stays in a unit, the lower are rents for a given level of housing service, even in markets without rent control. This “tenure discount” associated with longer stays is often motivated by a landlord’s desire to reduce turnover, avoid vacancy losses, and continue leasing to known tenants.<sup>6</sup> Consequently, a question should be included about the length of the family’s tenure in the unit.

A related factor that critically affects demand is the mobility of the household. This can vary enormously among countries. Strassman (1991) found that, in a given year, fewer than 5 percent of households in Colombo, Sri Lanka moved, whereas in Bangkok, Thailand about 20 percent of households moved in a year and in Seoul, Korea an astounding 43 percent moved. Including questions about the length of tenure in the survey can yield data that can be used to study such behavior. More elaborate housing questionnaires often add additional questions about previous residences and planned moves (see Mayo and others 1982 and Malpezzi 1994).

**SUPPLY.** As was discussed in the first section of this chapter, the supply of housing in any given country consists of the existing stock and of new construction. In any given year, well over 90 percent of the housing in a given market consists of the existing stock. Descriptive tabulations of housing characteristics, both on their own and cross-tabulated by relevant criteria such as income and tenure, can yield important insights into housing in the existing stock.

A more dynamic way to analyze the supply of housing from the existing stock is known as studying “filtering.” There are three ways of analyzing filtering (Green and Malpezzi 1997). The first way is to examine the incomes of the changing occupants of existing housing units over time and whether they “filter up” or “filter down” (Zais and Thibodeau 1983). The second way is to examine the price per unit of housing services for different parts of the housing stock—for example, low quality versus high quality housing (Lowry 1960). The third alternative is to examine how the quantity of the stock changes (Malpezzi, Ozanne, and Thibodeau 1987). For example, what effect does new construction have on the amount of low-quality housing? What are vacancy rates like at the bottom of the market? How fast do units depreciate? Each of these types of analysis can be done with data provided that a panel of data is collected. The answers to questions on rents and prices, household income, tenure, length of stay, housing characteristics, and age of the unit are key variables for filtering studies. The respondent should also be asked whether the household has had or currently has any plans to upgrade its dwelling.

**PROPERTY RIGHTS AND TENURE.** Another set of variables that needs to be collected in the questionnaire is

the set of variables related to tenure security. First and foremost, analysts need data on how long the household has lived in its current dwelling. Information on the type of rent control on the dwelling or any subsidy received by the household is often relevant for the study of tenure security since security is often related to these regulations. Other questions may need to be included in markets in which there is squatting or a mix of “traditional” and “formal” tenure.

Household surveys have a number of uses in studying property rights and tenure issues (Daniere 1992; Friedman, Jimenez, and Mayo 1988; Gyourko 1989; Jimenez 1984; Lim, Follain, and Renaud 1980). Questions relating to property rights and tenure should be drafted carefully to ensure that they reflect the current circumstances in the country of the survey. Thereafter, at a minimum, rights and tenure should be categorized in three ways: owning versus renting, informal versus formal/secure tenure, and public/social versus private ownership. These categories are often continua rather than mutually exclusive. For example, in Korea renting encompasses several payment systems, including periodic payment of rent, a deposit-based rental system (*chonsei*), and several mixed forms of deposit and periodic rent (*wolsei*). On the other hand, the British system of very long-term leases (99 years or more) is in some ways closer to owning than renting, even though periodic ground rent is paid and eventually the property reverts back to its residual owner.

**LAND AND INFRASTRUCTURE.** Since the provision of infrastructure is a core function of all governments, the proportion of households living on land served by basic infrastructure is of great interest to public policymakers. The benefits of the services can often be approximated by how they affect land value. The LSMS housing module should contain questions about the value of lots, as well as questions about their size, location, and the type of infrastructure to which they have access.

**HOUSING FINANCE.** Many of the questions relevant to housing finance are included in the savings and credit modules rather than the housing module (see Chapters 20 and 21). Of course the questions described in those chapters have to be tailored to local conditions. For example, the most common kind of mortgage in the United States (also found in many

other countries), a self-amortizing mortgage with a fixed interest rate and equal payments, can be completely described by four pieces of information: the interest rate, the loan amount, the loan term (duration), and any up-front fees. However, many other kinds of mortgages are possible. For example, interest rates may be tied to an index or payments and amortization schedules may vary (Buckley 1996; Chiquier and Renaud 1992).

Much can be learned from household survey data about how different kinds of households finance their housing and on what terms. Discrete choice models and cross-tabulations can be used to analyze these outcomes. Another finance issue that can be analyzed using household survey data is the relative inefficiency of “progressive building” (which is based on the stockpiling of materials and their use from time to time) compared to mortgage finance (see Renaud 1984).

In countries in which financing is subsidized for some borrowers or some kinds of households face very different finance terms than others, the value of different “deals” can be calculated in present value terms and then the distribution of these implicit transfers can be analyzed. World Bank (1989) demonstrated how to carry out a simple analysis of this type. Struyk and Turner (1986) demonstrated another way in which household survey data can be used to study the effects of finance on the housing market. They developed a simultaneous model of housing investment and demand for finance that can be used to test whether, and if so how much, finance availability affects housing investment.

**HOUSING AND EMPLOYMENT.** The importance of location with respect to workplace and other services was discussed above. When housing markets do not function well, this can prevent the efficient functioning of labor markets in general (Hughes and McCormick 1987; Johnes and Hyelak 1994; Mayo and Stein 1995). Another issue that must be tackled in some countries is the fact that in many specific enterprises, both public and private, employees’ housing is provided in conjunction with their employment. Enterprise housing in China is the most obvious example of this phenomenon, but company housing can often be found in noncommunist countries as well (Tolley 1991; Fishback 1992). For example, company housing is often associated with mining and other extractive industries when these are undertaken in remote areas.

If relevant, questions should be included in the housing module about employer- or enterprise-provided housing.

**MIGRATION.** Another issue that arises mainly in rural areas is the housing of migratory workers, such as itinerant agricultural laborers. This issue also sometimes arises in urban areas. For example, in China, the government classifies many urban households as “temporary.” This can make the choice of sampling frame particularly critical. Many obvious sampling frames, such as household registration lists, may systematically miss such households. Thus this kind of sample frame may need to be supplemented to ensure that these households are included in the sample.

**DATA FROM OTHER PARTS OF THE QUESTIONNAIRE.** Much housing analysis, especially studies of housing demand, relies on data gathered in other parts of the questionnaire. The main data needed for housing analysis from other parts of the questionnaire are summarized here so that survey designers will not overlook them.

It is reasonable to assume that the demand for housing is related to the household’s expectations about its long-term economic situation. Since housing consumption is related to long-run or permanent income, this suggests that permanent income rather than current income is the true determinant of housing consumption. Permanent income is, however, never directly observable and total household consumption is usually used to proxy for it (Hall 1978). Thus it is important for housing demand analysis that the questionnaire contain detailed consumption modules.

It is also useful for housing market analysts to have data on current income measures as well—for example, to analyze mortgage underwriting criteria or to study the targeting of housing subsidies. Because the qualification process for various subsidies and mortgage underwriting usually depends on current income rather than on permanent income or consumption, analysts need to know the household’s marginal propensity to consume out of its current income as well as its consumption. What would be even more useful for housing analysts would be a detailed analysis of the marginal propensity to consume housing out of different kinds of income (by type of employment, by the head of household versus the other household members, and so on). Thus these

types of data should also be collected in the relevant modules of the questionnaire.

Estimating patterns of demand requires data not only on prices and incomes but also on other determinants of demand such as the family’s preferences about housing, the family’s composition, and the household’s size (which is the most important single demographic variable affecting housing consumption). Other data that would be useful for analysis include the age of household head, the number of children in the household, and the sex of the head of the household. In some circumstances it may be appropriate to collect data on the household’s income, type of tenure, religion, or caste to use as proxies for taste.

#### *Survey Issues*

There are several important issues relating to the mechanics of implementing the housing module.

**SAMPLE.** Statistical methods are used to estimate the sample size required to answer a particular question to a desired degree of precision (Kish 1965). Experience suggests that roughly 500 observations are the minimum required from a given “housing market” (for example, a metropolitan area or a rural region) for useful analysis that cross-classifies data by tenure and other factors and that allows for nonresponses and other data problems. Because LSMS surveys tend to have national samples of 2,000–5,000 households, they are often unable to produce large enough subsamples in all but the largest metropolitan areas. This means that current LSMS designs are better suited to broad analyses of “national,” rural-urban, or regional housing markets. However, much research suggests that defining markets so broadly often obscures important differences among geographically disaggregated markets. Of course, resource constraints are a fact of life, and much can be done with surveys on the scale of the typical LSMS. Yet if housing market analysis is an important goal of an LSMS and if there appear to be different market conditions in different cities—or in different rural regions—in the country of the survey, serious consideration should be given to increasing the size of the sample or to over-sampling cities or regions of special interest. If the latter strategy is adopted, sample weights must be assigned to reflect this over-sampling.

**PANEL DATA.** Analyzing the dynamics of the housing market over time requires panel data. However, using

the household or the individual rather than the dwelling as the unit of observation can present comparability problems for housing analysts because households do not necessarily stay in the same dwelling between survey rounds. In previous LSMS surveys the housing unit has generally been used as the basis for the sample frame, which means that the survey followed the housing unit rather than its original occupants over time. While this has complicated analysis for some other issues, it is preferable for some housing analysis.

In some studies, such as the Mayo and others (1982) study of Egypt, retrospective questions were used as a proxy for a prior panel. Of course, this is not as good for analysis as proper panel data, as respondents often give inaccurate responses to retrospective questions because their memories of past events are imperfect.

Some key issues that need to be addressed when designing such a panel include the need to ensure that units that have dropped out of the stock are clearly coded to distinguish them from units that are temporarily unoccupied and the issue of how to bring newly constructed units into the panel over time. It must be possible to link each unit's data in one year's file to that in another year's file. It is essential to include a unique identifier code for each unit. Units that have been demolished, held vacant, or otherwise dropped out of the panel in the past should be identified, along with their current status. With regard to vacant units, survey designers should devise a short section of questions to be put to a respondent in a neighboring dwelling to discover, for example, how long the unit has been vacant, whether it is slated for demolition, and the rent at which it is being offered.

**COUNTRY-SPECIFIC QUESTIONS.** The need for survey designers to tailor the questionnaire carefully to local conditions cannot be overemphasized. For example, it is highly unlikely that bamboo would be used to construct houses in Moscow. It is just as important to tailor less obvious questions such as those about tenure and payment methods. See Malpezzi with Loux (1994) for examples of more detailed housing questionnaires.

### **The Housing Module**

This section introduces a draft housing module (presented in Volume 3) which, suitably modified, can be inserted into an LSMS questionnaire. "Suitably modi-

fied" deserves special emphasis. Every country is different in terms of the physical design of housing, its tenure, how it is paid for, and so on. The sample questionnaire introduced here should be considered only a starting point for designing an actual module. The initial design of any module should be thoroughly pretested to ensure that it is capable of yielding the required data. This sample questionnaire will not repeat questions that appear in other modules of the survey and are covered thoroughly in the relevant chapters. Note that this module contains a bare minimum of questions on water, sanitation and fuel use, which are suitable for describing basic living conditions and enumerating households' major expenditures on these items. If water, sanitation, or fuel use are of special interest in the survey, the questions in this draft module should be dropped, and the expanded submodules contained in the environment chapter (Chapter 14) should be inserted in their place.

Similarly, the draft module does not contain much on housing finance, since such questions are contained in the credit module introduced by Chapter 21 (and presented in Volume 3). If that module were to be dropped, some of the questions about credit for housing could be moved to this module. Additional questions can be found in the sample housing questionnaire in Malpezzi with Loux (1994).

The "long" draft module presented here is somewhat longer than that used in many past LSMS surveys. This is partly because it will support more analysis of housing market issues, rather than merely the description of living conditions and calculation of consumption of housing. It also includes water and sanitation questions that are suitable to situations in which households use multiple sources; includes questions on such transactions as deposits, "key money," and cooperative fees, which were rarely covered in previous LSMS surveys; and tries to cover the full range of housing market characteristics that exist in all regions of the developing world from Eastern Europe to Sub-Saharan Africa. In practice, only in very few countries will all of these additional questions need to be included in the module. In the places where a particular characteristic is rare, questions about that characteristic can be simplified or omitted. A shortened version of the questionnaire is presented after the main version to give an idea of how it can be shortened. In this case, some of the topics that allow study of housing market issues have been omitted, and the detail on living standards has been reduced. Again, the short version shown here is merely indica-

**Box 12.2 Cautionary Advice**

- *How Much of the Draft Module Is New and Unproven?* Almost all of the components of the draft housing module have been used either in past LSMS surveys or in special-purpose housing surveys.
- *How Well Has the Module Worked in the Past?* This module has been used for simple descriptive sketches of the housing conditions of the households, for which it has worked fairly well. One exception to this is that the modules included in past LSMS surveys have often included only one question on the household's source of water, which in many situations has not reflected the complexity of household water sources. Also, some of the housing cost questions have been ambiguous or insufficient. In particular, they have failed to make clear whether the rent includes utilities, and few surveys have included questions on any additional financial transactions such as key money or condominium or cooperative fees. However, while previous LSMS studies have made only limited use of the housing module, many other studies have been undertaken in developing countries that have made extensive use of such data. Mayo and others (1982) is probably the best single example.
- *Which Parts of the Module Most Need to Be Customized?* A great deal of the module needs to be carefully customized to reflect the housing conditions in the country where the survey is to be fielded. Many aspects of housing vary greatly from country to country, including the predominant types of dwellings, the materials that they are made of, the kinds of amenities that are indicators of living standards, and the form in which different housing-related expenditures are made. For example, questions on privatization of state-owned dwellings, on how well elevators operate, and on the adequacy and costs of heating will be relevant in surveys in Eastern European countries but not in countries in Sub-Saharan Africa.

tive; exactly which subset of questions should remain in a shortened version will depend on the circumstances. For example, the short version shown here omits questions on key money and other such deposits, but if they are relevant to a country setting, they should be included, even in a short version of the questionnaire.

**Notes on the Housing Module**

This section briefly discusses the definition of key concepts and other specific points in the module, following the numbering system of the longer version of

the module. When the module is going to be used in an actual LSMS survey, it is important to produce a manual that includes a more detailed checklist of definitions both for survey workers and for future users of the data. The U.S. Census report on the American Housing Survey 1995 (which can be downloaded from [www.census.gov](http://www.census.gov)) provides a general example of such documentation. See also Malpezzi, Bamberger, and Mayo (1982) and Malpezzi (1994) for further examples.

For housing analysts to be able to use the housing data from an LSMS survey, the survey must also yield accurate, reliable information on related topics, such as household size and composition and household income. It is assumed in this chapter that these key collateral data are indeed collected in accordance with the discussion in the other chapters in this book.

It cannot be emphasized enough that survey designers will need to revise and pretest the questionnaire to bring it in line with local conditions. For example, there are not very many houses in Cracow, Poland that have felt walls or thatched roofs, and detailed questions about heating systems will be irrelevant in Accra, Ghana. While this section does not address the issue of country-specific relevance with regard to every question, survey planners should do so themselves when they are designing the questionnaire for their particular survey.

**Part A: Description of the Dwelling**

Part A of the housing module is designed to yield data that give a basic description of the dwelling.

Question A1 asks whether the dwelling is the household's primary residence and, if it is not, redirects the interview to be about the primary residence. For measuring living standards, it is most important to know about the conditions of the primary residence since those are the ones that pertain to the household most of the time and best reflect the quality of infrastructure available to the household. If the survey's purpose were only analysis of housing markets, gathering information about the costs and quality of secondary residences would be a perfectly reasonable option.

There are at least three ways to deal with secondary residences. In past LSMS surveys the issue was completely ignored, and Question A1 was not used. Although this is not technically correct, no complaints have ever been made to the central LSMS team on the

subject. One reason this issue has been ignored is that in most countries where LSMS surveys have been done, secondary residences are rare and pertain only to the extreme upper end of the welfare distribution. Moreover, the richest frequently have the highest non-response rates and even when they do respond, their expenditure, income and wealth are probably underestimated since LSMS questionnaires are designed to be applicable to the broad range of society with special emphasis on the poor. Thus ignoring this issue in the past may not have had much empirical impact on most of the analysis done with the data.

A second option is to use just the simple question included here. It will give some information on how important the topic is in the country, and will allow sampling weights to be adjusted. A third option is to deal with the issue of secondary residence much more fully. This will be appropriate where secondary residences are relatively common and their ownership extends to a wider range of society (for example, in Finland, where about 20 percent of households have a secondary dwelling). To deal with the issue fully will mean not only directing the interview on housing quality to the primary residence, but also adding questions about at least current expenditures on the secondary dwelling, and probably adding questions on its value as an asset. Whatever approach is taken in the questionnaire should accord with how secondary units are treated in the sample. Are they included or excluded from the sample frame? Are they substituted out if detected during interviewing? Are the sampling weights adjusted for households that own or geographic areas that contain secondary residences?

Each person has a commonsense notion of what is meant by such terms as “house,” “household,” “room,” and so on, but these notions may differ from person to person. For example, is a “bathroom” also counted as a “room”? Accurate use of survey data is only possible if such definitions are consistent—in other words, if all of the survey interviewers have the same definition of each concept. For this reason, some definitions of common but important housing concepts are presented here. Many of the sample definitions will have to be modified to suit country conditions.

For example, consider Question A11 on rooms. The definition of “room” will vary from country to country. A sample definition that can be used as a starting point, adapted from the U.S. Census definition, is: “whole rooms used for living purposes, such as living

rooms, dining rooms, bedrooms, finished attic or basement rooms, recreation rooms, permanently enclosed porches suitable for year-round use, lodger’s rooms, and rooms used as offices or for business purposes. A divided room is separate if there is a partition from floor to ceiling but not if the partition is impermanent or made only of shelves or cabinets. Not included are bathrooms, halls, vestibules, balconies, alcoves, closets, unfinished attics, or basements, unenclosed porches. If a room is used by occupants of more than one unit, the room is included with the unit from which it is most easily reached.”

Separate questions should be developed for particular types of rooms or structural features that are especially important in the country surveyed. For example, many questionnaires ask how many bedrooms a unit has. In the Ghanaian survey that was analyzed in Malpezzi, Tipple, and Willis (1990), separate questions were asked about unenclosed verandahs, because households with this feature tend to make considerable use of it. It is not important that there is double-counting in this case, since a bedroom would be counted both as a room and a bedroom. What is important is that the special rooms are either always double counted or never double counted and that the documentation makes clear which is the case.

**DWELLING UNIT.** A dwelling is an accommodation unit that contains one or more households. It may be a detached house, a villa, part of a flat, a shack, a tent, a separate room, or a houseboat. There may be several dwellings in a structure.

**STRUCTURE.** A structure is a physically separate entity such as a house, an apartment building, or a tent. It may contain one or more dwelling units.

**BEDROOMS.** The number of bedrooms in the unit is the sum total of all separate rooms that are used regularly for sleeping, even if they are also used for other purposes. Rooms reserved for guests’ sleeping are counted as bedrooms. On the other hand, rooms used regularly for other purposes, even though used occasionally for sleeping, are not counted as bedrooms. All bedrooms are also counted as rooms.

Question A14 asks about the area of the unit. In some countries, such as Korea, households are likely to know this area precisely. In other countries they will only be able to produce a rough estimate.

A number of questions in this and other sections are questions for which households may have only approximate answers. In some cases, such as the area of the unit, an alternative approach is possible; for example, if there is enough interview time, it may be possible for the interviewer actually to measure the dwelling unit. For other questions, such as the age of the dwelling, no such alternative may exist.

Generally, it is better to get an approximate answer to the right question than a precise answer to a useless question. This may seem obvious, but census bureaus around the world mistakenly exclude important questions because they are likely to be measured with error. It is certainly important to understand the consequences of such errors—in particular to understand the difference between biased estimates and imprecise estimates. For example, studies have shown that households tend to give answers to questions about the age of their unit that contain a significant degree of error. However, if they are as likely to overestimate as to underestimate, the statistics based on this data (such as the mean age of dwellings of a certain type) will be unbiased, although these estimates will be less precise than if the respondents had a very good idea of the age of the unit. A further discussion of this issue can be found in Follain and Malpezzi (1981).

### **Part B: Housing Services**

If water, sanitation, or fuel use are of special interest in a given survey, the expanded modules contained in Chapter 14 on the environment will be better starting points for questions on those subjects than the questions given here. The questions on such topics included in this housing module can only yield descriptive information. If the specialized modules contained in the environment chapter are used, these questions should be omitted. It would be natural to put the housing module next to the water and sanitation modules in the questionnaire and possibly next to the fuel module as well (though this might just as logically be placed next to the consumption module).

The questions on water sources included in this draft housing module distinguish between rainy and dry seasons. In some countries this distinction can be omitted. The module also distinguishes sources depending on whether they are used for drinking and cooking or for bathing and washing. In a few countries that are highly urbanized and have very well-developed water systems, this distinction can be

dropped. However, households in rural areas probably use different sources for these two purposes until there is a considerable amount of infrastructure in their areas, which means that this distinction is pertinent in most countries.

The questions about “what is the main source of water...” are a little tricky to word. These questions aim to yield data on the type of access that a household has, not on the body of water that feeds into the central pipeline. Thus great care should be taken in translating these questions. Similarly, there are many different possible sources, and they can be called different things in different places (for example, a standpipe versus a public tap). The basic idea is to devise use answer codes that convey something about the likely safety and convenience of each source, without devising so many codes as to overwhelm the interviewer or the respondent.

Sanitation systems (flush toilets, pit latrines, bucket systems, and so on) are another example of something that varies tremendously from country to country. A housing unit is classified as having a bathroom if it has a room attached to the house with at least one of the following: a toilet, and a bathtub, a shower, or a sink with running water. If a unit has these facilities but the toilet and at least some washing facilities are not in the same room, then the unit does not have a bathroom.

A kitchen is a room set aside for preparing food. It must have a stove or other facility for cooking and may have a sink and a refrigerator or icebox as well. A complete kitchen has all three facilities. A kitchen is also counted as a room if it is enclosed.

### **Part C: Dwelling Expenditures**

Part C of the module focuses on household expenditures on housing. Obviously, how these questions are asked will vary from place to place. In particular, questions about expenditure are inextricably bound up with questions about the form of housing tenure, and this varies from place to place. Often, units are either owned outright or rented, but there are many other forms of tenure in some countries. In Korea many households have a form of tenure called *chonshei*, which is similar to renting but, instead of paying periodic rent, the household puts down a large refundable lump sum as a deposit, often as much as half of the value of the unit. Other forms of tenure in Korea include owning outright and renting, but there are

also mixed forms, such as households that put down a smaller deposit and then pay a periodic rent, *wolsei*.

This section is closely related to the chapter on credit, which introduces a draft credit module in which data are collected on mortgage transactions (see Chapter 21).

In addition to collecting accurate, reliable data on expenditures associated with housing, it is extremely important to get some sense of whether these particular households are facing market prices and engaged in arms-length transactions. For example, it is important to design the questionnaire to find out whether the government provides a household with its dwelling. In that case, analysts might want to know what the rent is for other purposes than as an indication of the state of the market. If a household is renting its dwelling from a close relative, the household may be paying a lower-than-market rent. In some cultures being a member of a kinship group implies that the household gets a discount. If this information is collected in the survey, analysts can study the size of these discounts.

Questions about payments that households make for their utilities are in this section, and use a recall period of the previous month. This should work well in places where most of these items are billed for on a monthly basis. In places where this is not the case, it may be preferable to ask respondents about some of these items earlier in the module when the amenity is discussed. For example, questions about different forms of payments for the different sources of water can be interwoven into that section. Chapter 14 on environment covers the most detailed set of water charges, and differentiates many of the questions according to the type of source and the different ways in which charges for it may be made, illustrating this idea of interweaving expenditure with use and amenity questions. Expenditures on fuel can be included in the housing module, in the consumption module, or in a specialized fuel use module, with increasing detail possible in each case. Naturally they need be put in only one of those places, though this book illustrates their placement in each of the three.

Some households rent out part of their dwelling. It is important to calculate the net costs (payments out minus rent coming in) of the dwelling. For analysis of crowding, it may be useful to get further information about the number of rooms rented out and the number of persons who occupy them. It is important to

make clear for the interviewer and for the respondent which rental income is covered here and which is covered in the transfers and other nonlabor income chapter. Only data on the rental income from the dwelling to which the interview pertains are captured in the draft housing module. Income from the rental of other dwellings where the respondent does not live is covered in the transfers and other nonlabor income module introduced by Chapter 11 rather than in the housing module.

#### ***Part D: Household Opinions About Their House and Neighborhood***

The purpose of this section is to identify the aspects of the house and neighborhood with which households are most and least satisfied. Only a few general opinion questions have been included in the draft housing module about households' satisfaction with their unit and their neighborhood. Hedonic price studies of the United States suggest that such general opinions are closely associated with housing prices but that once such general questions have been asked, more detailed questions (for example, about households' satisfaction with schools, public safety, and so on) are not generally statistically significant.<sup>7</sup>

However, there may be situations in which it is worthwhile to ask additional questions about housing and neighborhood satisfaction. For example, it is plausible that different neighborhood characteristics may be valued differently in different countries. For example, consider a country with a highly stratified educational system, where attending primary school in a particular location leads to the opportunity to attend a prestigious secondary school and university. The value of this may be capitalized into housing prices and may be highly significant in such a country.

If the list of neighborhood questions is expanded, Malpezzi with Loux (1994) and especially the American Housing Survey have many examples of potential questions. It is possible either to leave open the list of aspects with which they are satisfied or dissatisfied and then to post-code them or to include a list in the questionnaire on the basis of a pilot survey.

#### ***Part E: Planned Moves and Upgrades***

A household can easily change its consumption of food either up or down by purchasing more or less food in a particular day or week. Changing a household's consumption of housing is more difficult and

costly. The household must either move or upgrade the unit in which it already lives.

Since households move so infrequently and this moving process is fundamental to understanding the state of the housing market, it is sometimes useful to ask retrospective questions about the previous unit in which the household lived or prospective questions if the household is planning to move. The usefulness of these questions and the way they are worded will vary from place to place. In countries like Korea, households move on average every two years, whereas in other countries such as Egypt households may move as infrequently as every 15 years. Also, people in different cultures have different attitudes about prospective questions.

#### *Housing-Related Questions in Other Modules*

The strong links between housing finance questions and the credit module, the transfers and other non-labor income module, and the specialized water, sanitation and fuel submodules contained in the environment module have already been noted. It has also been noted that having accurate, reliable information on income, household size and composition, and commuting from other modules in the survey is important for housing analysis. Information on housing costs can be gathered in the community questionnaire.

Questions about household composition in the roster module should be drafted in such a way as to distinguish temporary accommodation from permanent accommodation. When units are shared by more than one household, this should be clearly indicated. In some countries it is important to indicate whether the landlord lives in the building or to clarify kinship relations between households in the unit. In addition, because many developing countries have surprisingly high vacancy rates, at least in some parts of the market (Mayo and others 1982; Struyk 1988), when this is of particular interest, it can be useful to devise questions that yield the data necessary for analysts to study the extent and incidence of vacancies and their determinants. The identification and control page (see Chapter 4 on metadata) has a question showing why interviews could not take place in the selected dwelling. One of the response codes is that the dwelling was vacant. It would be possible to add follow-up questions there that would be asked of neighbors to gather some information on vacancies.

For many purposes it is useful—and sometimes very important—to include information about the

location of the unit within a city or other geographical unit. These variables can be coded from the unit's address. For example, distances to the center of the city should be coded for urban units. See Ingram (1984) and Mohan (1994) for examples of the use of locational data in housing market analysis.

### **Appendix 12.1 What is the “Price” of Housing?**

There is a difference between the way in which economists use the term “price” and the way in which housing analysts, real estate professionals, and other noneconomists often use the term.

Economists generally define rent, the periodic expenditure for housing, as the product of the price per unit of housing,  $P$ , and the real estate services yielded by the unit,  $Q$ . Thus  $R = PQ$ . Rent and this associated price per unit of service,  $P$ , are “flow” (per period) concepts. The physical real estate itself is durable, so  $Q$  is a “stock” concept. A stock (housing asset) yields a flow of services over time.

Of course, many readers will know that the flow “rent” can be translated into the stock concept of “value”:  $V = R/i$ , where  $i$  is the capitalization rate. Housing value, the stock analog to rent, is also known to economists as the asset price of the unit. When real estate brokers and others use the term house price, or unit price, they are referring to this present value measure or asset price,  $V$ , rather than the flow price per unit of service,  $P$ , as described above. When economists use the term price, they are often referring to  $P$ . However, even economists sometimes loosely refer to  $V$  as price, although careful economists will usually use the term “asset price.” In any event, the context should make the distinction clear.

Note that if, by the assumptions of their model or analysis, analysts standardize the quantity of real estate services produced (say, in square feet of a given homogenous level of quality, including location), then rent and flow price are basically synonymous. More precisely, rent and flow price  $P$  are proportional, since by assumption  $Q$  is fixed.

Housing economists use a number of different methods to construct indexes of the price of housing. The main types of methods: simple medians and averages, Laspeyres, Paasche, Divisia, and related time series indexes, hedonic price indexes, repeat sales indexes, user cost models, and hybrid methods. These methods are described briefly in the following paragraphs;

Malpezzi and Green (1998) provides a more detailed discussion.

#### *Simple Medians or Averages*

The most commonly used measures of this type in the United States are median sales prices for existing housing (which are published by the National Association of Realtors) and Census Median House Prices (values for owner occupiers and rents for renters). The method is, in general, self-explanatory. A big advantage of this type of measure is its simplicity and the fact that it allows rough comparisons over time and across markets. The biggest disadvantage is that this type of measure does not usually control for differences in the quantity of housing services,  $Q$ , across markets or over time.

A number of studies suggest that, while these simple indexes are not adjusted for quality differences, quantity generally varies less than price in such a sample. Thus the studies conclude that these simple measures, while imperfect, do include valuable price information.

#### *The Laspeyres Price Index and Related Indexes*

Familiar examples include consumer price indexes and implicit price deflators from national income accounts, which are available in virtually all countries. These are generally constructed by taking a sample of units in some base year and revisiting the units over time, appraising them, and computing any percentage changes. The familiar Laspeyres indexes are constructed as:

$$I_t = (P_t Q_0 / P_0 Q_0) 100$$

where  $I$  is the index,  $P$  is the price per unit of housing services,  $Q$  is the quantity of housing, and subscripts denote time. Time 0 is the base year or period, and time  $t$  is any year, forward (positive  $t$ ) or backward (negative  $t$ ). Thus the index is the ratio of what is spent in time  $t$  to what is spent in time 0, holding what is purchased constant to the “bundle” purchased in time 0.

Related indexes, including Paasche and chain indexes, are discussed in Afriat (1977) and Diewert (1991). Indexes differ in how the bundle is fixed or varied. The U.S. Department of Commerce has recently moved from Laspeyres to chain indexes (with a constantly changing bundle) for most time series.

Laspeyres and related indexes have much to recommend them, but they do have some disadvantages.

Generally, these are time-series indices only. That is, if there is one housing consumer price index for, say, Monterey, and another housing consumer price index for Tijuana, it is possible to compare how fast prices are rising in the two cities but not to discover which city is more expensive. Also, the results may vary depending on which “bundle” (typical housing unit) is chosen. Ideally, analysts would like to hold the bundle fixed, but as prices change over time, the typical bundle consumed changes in real life, even if not in the index.

#### *Hedonic Indexes*

These are constructed by regressing rent or value against characteristics of the unit and its location. Then analysts use the coefficients to predict rent or value for “standard” units. Most often these are done for one point in time, but they can be done over time as well. These methods have good theoretical and intuitive foundations and are discussed in detail in Malpezzi, Ozanne, and Thibodeau (1980). However, they involve substantial data requirements and analytical work.

#### *Repeat Sales Indexes*

These indexes are constructed after surveying units that have been sold twice. Although they are constructed using regression methods, intuitively these indexes are roughly similar to annualizing and averaging the percentage growth in sales prices over time. These indexes are time-series only. They have the advantage of being based on actual transaction prices, but most units are not sold in any given period, so using repeat sales misses a lot of information. Also, units that sell are not necessarily representative of all units, and sometimes it can be hard to tell whether  $Q$  for a unit has changed (for example, due to remodeling). Repeat sales indexes are thoroughly discussed in Wang and Zorn (1997).

#### *The User Cost Method*

The idea behind this method is simple: it calculates what a “user” of the house really pays (or would pay) net of financing, taxes, maintenance, inflation, and so on. These measures are generally time-series (for example, Hendershott and Shilling 1982) but can be done for one point in time (for example, Follain 1982). The user cost method incorporates a model of what actually determines prices, and it accounts for the effects of taxation, inflation, and maintenance on prices.

### Hybrid Indexes

These indexes combine (usually) two of the above methods. Hybrid indexes can be time-series, for one point in time, or both. For example, hedonic and repeat sales methods can be combined (as in Case and Quigley 1991) as can hedonic and user cost methods (as in Follain 1982).

### Notes

The author is indebted to Margaret Grosh, Paul Glewwe, and Fiona Mackintosh for comments on previous versions.

1. Specifically, the income elasticity of demand is the percentage change in expenditure given a percentage change in income. See Meier (1983) for elaboration.

2. Many existing LSMS surveys have collected substantial housing information that has not yet been used in analysis.

3. For example, very few empirical analyses have been done of U.S. rural housing markets despite the vast literature in that country. See Vandell (1997) for a review and discussion.

4. An hedonic index is a regression of rents (or house values) against the characteristics of the units. See Appendix 12.1 for a more detailed explanation.

5. In the case of the owner-occupied imputation, the question must be asked in such a way that the respondent assumes that this would be such an arms-length transaction.

6. See Malpezzi, Ozanne, and Thibodeau 1980, pp. 78–9.

7. The lack of significance does not prove or demonstrate that specific things like schools or public safety do not matter but rather that, once general neighborhood and unit satisfaction are taken into account, additional specific questions do not seem to add much information.

### References

- Afriat, S. 1977. *The Price Index*. New York: Cambridge University Press.
- Angel, Shlomo, and others. 1986. "The Land and Housing Markets of Bangkok: Strategies for Public Sector Participation." Planning and Development Cooperative, Thailand National Housing Authority and the Asian Development Bank, Bangkok.
- Bertaud, Alain, and Bertrand Renaud. 1994. *Cities Without Land Markets: Lessons of the Failed Socialist Experiment*. World Bank Discussion Paper 227. Washington, D.C.
- Betancur, John J. 1987. "Spontaneous Settlement Housing in Latin America: A Critical Examination." *Environment and Behavior* 19: 286–310.
- Bramley, Glen. 1993. "The Impact of Land Use Planning and Tax Subsidies on the Supply and Price of Housing in Britain." *Urban Studies* 30 (1): 5–30.
- Bromley, Daniel W. 1989. "Property Relations and Economic Development: The Other Land Reform." *World Development* 17: 867–77.
- Buckley, Robert M. 1996. *Housing Finance in Developing Countries*. New York: Oxford University Press.
- Burns, Leland S., and Leo Grebler. 1977. *The Housing of Nations: Advice and Policy in a Comparative Framework*. London: Macmillan.
- Case, Bradford, and John M. Quigley. 1991. "The Dynamics of Real Estate Prices." *Review of Economics and Statistics* 22 (1): 50–58.
- Chiquier, Loic, and Bertrand Renaud. 1992. "Alternative Mortgage Instruments in Distorted Housing Systems: How Useful is the Dual Rate Adjustable Mortgage?" World Bank, Industry and Mining Division, Energy, Mining, and Telecommunications Department, Washington, D.C.
- Daniere, Amrita. 1992. "Determinants of Tenure Choice in the Third World." *Journal of Housing Economics* 2 (2): 159–84.
- Deaton, Angus, and John Muellbauer. 1980. *Economics and Consumer Behavior*. New York: Cambridge University Press.
- Dillinger, William. 1991. *Urban Property Tax Reform*. Washington, D.C.: World Bank, United Nations Centre for Human Settlements (Habitat), and United Nations Development Programme.
- Diewert, W. E., ed. 1991. *Price Level Measurement*. New York: North Holland.
- Dowall, David. 1991. *The Land Market Assessment: A New Tool for Urban Management*. Washington, D.C.: World Bank, United Nations Centre for Human Settlements (Habitat), and United Nations Development Programme.
- Farvaque, Catherine, and Patrick McAuslan. 1992. "Urban Land Policies and Institutions in Developing Countries." Urban Management Program Paper UMP-5. World Bank and United Nations Development Programme, Washington, D.C.
- Ferchiou, Ridha. 1982. "The Indirect Effects of New Housing Construction in Developing Countries." *Urban Studies* 19: 167–76.
- Fishback, P.V. 1992. "The Economics of Company Housing." *Journal of Law, Economics and Organization* 8 (2): 346–65.
- Follain, James R. 1982. "Does Inflation Affect Real Behavior? The Case of Housing." *Southern Economic Journal* 48 (3): 570–82.
- Follain, James R., and Stephen Malpezzi. 1981. "Are Occupants Accurate Appraisers?" *Review of Public Data Use* 9 (1): 47–55.
- Follain, James R., and Emmanuel Jimenez. 1985a. "The Demand for Housing Characteristics in Developing Countries." *Urban Studies* 22.
- . 1985b. "Estimating the Demand for Housing Characteristics: A Survey and Critique." *Regional Science and Urban Economics* 15 (5): 421–32.

- Follain, James R., Gill-Chin Lim, and Bertrand Renaud. 1980. "The Demand for Housing in Developing Countries: The Case of Korea." *Journal of Urban Economics* 7 (3): 315-36.
- Freeman, A. Myrick. 1979. *The Benefits of Environmental Improvement: Theory and Practice*. Washington, D.C.: Resources for the Future.
- Friedman, Joseph, Emmanuel Jimenez, and Stephen K. Mayo. 1988. "The Demand for Tenure Security in Developing Countries." *Journal of Urban Economics* 29 (2): 185-98.
- Gackenheimer, Ralph, and Carlos Henrique Jorge Brando. 1987. "Infrastructure Standards." In Lloyd Rodwin, ed., *Shelter Settlement and Development*. Boston, Mass.: Allen and Unwin.
- Gilbert, Alan. 1989. *Housing and Land in Urban Mexico*. University of California, Center for US-Mexican Studies, San Diego.
- Green, Richard K., and Stephen Malpezzi. 1998. "A Primer on U.S. Housing Markets and Policies." University of Wisconsin, Center for Urban Land Economics Research Working Paper, Madison, Wis.
- Gyourko, Joseph, and Jaehye Kim Han. 1989. "Housing Wealth, Housing Finance, and Tenure in Korea." *Regional Science and Urban Economics* 19 (2): 211-34.
- Hall, Robert E. 1978. "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence." *Journal of Political Economy* 86 (December): 971-87.
- Hendershott, Patric H., and James D. Shilling. 1982. "The Economics of Tenure Choice, 1955-1979." In C. F. Sirmans, ed., *Research on Real Estate*. Greenwich, Conn.: JAI Press.
- Hughes, G. A., and B. McCormick. 1987. "Housing Markets, Unemployment, and Labor Market Flexibility in the U.K." *European Economic Review* 31 (3): 615-45.
- Ingram, Gregory K. 1984. "Housing Demand in the Developing Metropolis: Estimates from Bogota and Cali, Colombia." World Bank Staff Working Paper 633, Washington, D.C.
- Ingram, Gregory K., and Alan Carroll. 1981. "The Spatial Structure of Latin American Cities." *Journal of Urban Economics* 9 (2): 257-73.
- Jaffe, Austin J. 1993. "Property Rights and Market Behavior in Eastern European Housing Reforms." Paper presented at the American Real Estate and Urban Economics Association meeting, January, Anaheim, Cal.
- Jaffe, Austin J., and Demetrios Louziotis Jr. 1996. "Property Rights and Economic Efficiency: A Survey of Institutional Factors." *Journal of Real Estate Literature* 4 (2): 137-59.
- Jimenez, Emmanuel. 1982a. "The Economics of Self Help Housing: Theory and Some Evidence from a Developing Country." *Journal of Urban Economics* 11 (2): 205-28.
- . 1982b. "The Value of Squatter Dwellings in Developing Countries." *Economic Development and Cultural Change* 30 (4): 739-52.
- . 1984. "Tenure Security and Urban Squatting." *Review of Economics and Statistics* 66 (4): 556-67.
- Johnson, Thomas E. Jr. 1987. "Upward Filtering of the Housing Stock." *Habitat International* 11 (1): 172-94.
- Johnes, G., and T. Hyclak. 1994. "House Prices, Migration and Regional Labor Markets." *Journal of Housing Economics* 3 (4): 312-29.
- Kiamba, Makau. 1989. "The Introduction and Evolution of Private Landed Property in Kenya." *Development and Change* 20 (1): 121-48.
- Kim, Jeong-Ho. 1991. "Housing Program Evaluation Using the Present Value Model: A Korean Experience." Paper presented at a World Bank Seminar on Korean Housing Markets and Policy, June, Washington, D.C.
- Kish, Leslie. 1965. *Survey Sampling*. New York: Wiley.
- Lee, Kyu Sik. 1992. "Spatial Policy and Infrastructure Constraints on Industrial Growth in Thailand." *Review of Urban and Regional Development Studies* 4.
- Lee, Kyu Sik, and Alex Anas. 1992. "Costs of Deficient Infrastructure: The Case of Nigerian Manufacturing." *Urban Studies* 29 (7): 1071-92.
- Lim, Gill-Chin, James R. Follain, and Bertrand Renaud. 1980. "The Determinants of Homeownership in a Developing Economy." *Urban Studies* 17 (1): 13-23.
- Lowry, Ira S. 1960. "Filtering and Housing Standards." *Land Economics* 35: 362-70.
- Malpezzi, Stephen. 1984. "Analyzing an Urban Housing Survey." Infrastructure and Urban Development Department Discussion Paper UDD-52. World Bank, Washington, D.C.
- . 1994. "Getting the Incentives Right: A Reply to Robert-Jan Baken and Jan Van Der Linden." *Third World Planning Review* 16 (4): 451-66.
- . 1996. "Notes on Consumer's Surplus." University of Wisconsin, Department of Real Estate and Urban Land Economy, Madison, Wis.
- . 1998. "Welfare Analysis of Rent Control With Side Payments: A Natural Experiment in Cairo, Egypt." *Regional Science and Urban Economics* 28 (6): 773-75.
- . 1999. "Economic Analysis of Housing Markets in Developing and Transition Economies." In Paul Chesire and Edwin S. Mills, eds., *Handbook of Regional and Urban Economics*. New York: North Holland.
- Malpezzi, Stephen, with Suzanne Loux. 1994. "Two Stylized Housing Questionnaires." Center for Urban Land Economics Research Working Paper. University of Wisconsin, Madison, Wis.
- Malpezzi, Stephen, and Stephen K. Mayo. 1987a. "The Demand for Housing in Developing Countries." *Economic Development and Cultural Change* 35 (4): 687-721.

- . 1987b. "User Cost and Housing Tenure in Developing Countries." *Journal of Development Economics* 25: 197–220.
- . 1997. "Getting Housing Incentives Right: A Case Study of the Effects of Regulation, Taxes and Subsidies on Housing Supply in Malaysia." *Land Economics* 73 (3): 372–91.
- Malpezzi, Stephen, Michael Bamberger, and Stephen K. Mayo. 1982. "Planning an Urban Housing Survey." Infrastructure and Urban Development Department Discussion Paper UDD-42. World Bank, Washington, D.C.
- Malpezzi, Stephen, Larry Ozanne, and Thomas Thibodeau. 1980. *Characteristic Prices of Housing in 59 SMSAs*. Washington, D.C.: The Urban Institute.
- . 1987. "Microeconomic Estimates of Housing Depreciation." *Land Economics* 63 (4): 373–85.
- Malpezzi, Stephen, Graham Tittle, and Kenneth Willis. 1990. *Costs and Benefits of Rent Control: A Case Study in Kumasi, Ghana*. World Bank Discussion Paper 74. Washington, D.C.: World Bank.
- Mayo, Stephen K. 1986. "Sources of Inefficiency in Subsidized Housing Programs: A Comparison of U.S. and German Experience." *Journal of Urban Economics* 20: 229–49.
- Mayo, Stephen K., and James Stein. 1995. "Housing and Labor Market Distortions in Poland: Linkages and Policy Implications." *Journal of Housing Economics* 4: 153–82.
- Mayo, Stephen K., and others. 1982. *Informal Housing in Egypt*. Cambridge, Mass.: Abt Associates.
- Meier, Gerald M., ed. 1993. *Pricing Policy for Development Management*. Baltimore, Md.: Johns Hopkins University Press.
- Mohan, Rakesh. 1994. *Understanding the Developing Country Metropolis: Lessons from the City Study of Bogota and Cali, Colombia*. New York: Oxford University Press.
- Olsen, Edgar O. 1987. "The Demand and Supply of Housing Services: A Critical Review of the Empirical Literature." In E.S. Mills, ed., *Handbook of Regional and Urban Economics*. Amsterdam: Elsevier.
- Ozanne, Larry, and Raymond J. Struyk. 1978. "The Price Elasticity of Supply of Housing Services." In L. S. Bourne and J. R. Hitchcocks, eds., *Urban Housing Markets: Recent Directions in Research and Policy*. Toronto: University of Toronto Press.
- Pejovich, Svetozar. 1990. *The Economics of Property Rights: Towards a Theory of Comparative Systems*. Boston, Mass.: Kluwer.
- Rakodi, Carole. 1987. "Upgrading in Chawama, Lusaka: Displacement or Differentiation?" *Urban Studies* 25 (4): 791–811.
- Renaud, Bertrand. 1980. "Resource Allocation to Housing Investment: Comments and Further Results." *Economic Development and Cultural Change* 28 (2): 189–99.
- . 1984. "Housing and Financial Institutions in Developing Countries." Staff Working Paper 658. World Bank, Washington, D.C.
- Sanyal, Biswapriya. 1981. "Who Gets What, Where, Why and How: A Critical Look at Housing Subsidies in Zambia." *Development and Change* 12: 409–40.
- Strassman, W. Paul. 1982. *The Transformation of Urban Housing: The Experience of Upgrading in Cartagena*. Baltimore, Md.: Johns Hopkins University Press.
- . 1991. "Housing Market Interventions and Mobility: An International Comparison." *Urban Studies* 28 (5): 759–71.
- Struyk, Raymond J. 1982. "Upgrading Existing Dwellings: An Element in the Housing Strategies of Developing Countries." *Journal of Developing Areas* 17: 69–76.
- . 1988. "Understanding High Vacancy Rates in a Developing Country: Jordan." *Journal of Developing Areas* 22 (3): 373–80.
- Struyk, Raymond, and Margery Turner. 1986. *Housing Finance and Quality in Two Developing Countries*. Washington, D.C.: Urban Institute Press.
- Tittle, A. Graham, and Kenneth G. Willis, eds. 1991a. *Housing the Poor in the Developing World: Methods of Analysis, Case Studies and Policy*. London: Routledge.
- . 1991b. "Tenure Choice in Kumasi, Ghana." *Third World Planning Review* 13 (1): 27–45.
- Tolley, George S. 1991. *Urban Housing Reform in China*. Washington, D.C.: World Bank.
- Vandell, Kerry D. 1997. "Improving Secondary Markets in Rural America." Federal Reserve Bank of Kansas City, Kansas City.
- Wang, Ferdinand T., and Peter M. Zorn. 1997. "Estimating House Price Growth with Repeat Sales Data: What's the Aim of the Game?" *Journal of Housing Economics* 6: 93–118.
- World Bank. 1989. "Malaysia: The Housing Sector; Getting the Incentives Right." World Bank Sector Report 7292-MA, Washington, D.C.
- . 1993a. "Housing: Enabling Markets to Work." World Bank Policy Paper, Washington, D.C.
- . 1993b. "Russia Housing Reform and Privatization: Strategy and Transition Issues." Report 11868-RU, Washington, D.C.
- Yu, Fu-Lai, and Si-Ming Li. 1985. "The Welfare Cost of Hong Kong's Public Housing Program." *Urban Studies* 2: 133–140.
- Zais, James, and Thomas Thibodeau. 1983. *The Elderly and Urban Housing*. Washington, D.C.: The Urban Institute.

# 13

## Community and Price Data

*Elizabeth Frankenberg*

Multitopic household surveys like the LSMS are designed to gather data to be used for analyzing household welfare, including analysis of access to and use of social services, the effects of government policies on living conditions, and how households behave in response to changes in the economic environment or in government programs. To meet these objectives, multitopic household surveys often collect data not only at the household level but also at the community level and even, in some cases, at the level of facilities (such as health clinics or schools).

Collecting community or facility-level data is desirable for two reasons. First, the government programs and services that affect individuals are often implemented and provided at the community level. Thus household surveys that collect information at both the household and the community level yield more policy-relevant data than those that only collect household data. Second, it is efficient to collect information about the shared environment in which households operate from community leaders or from community members who are particularly knowledgeable about key subjects rather than from each household individually.

In LSMS surveys, these additional data are collected in a community questionnaire that is administered separately from the household questionnaire. The informants (called “informants” to distinguish them from the “respondents” to the household questionnaire) selected for these community questionnaires vary depending on the specific objectives of the survey, but they can include community members, market traders, and staff at relevant facilities and institutions, such as nurses and teachers.

The community-level information collected in LSMS surveys typically includes information on infrastructure, employment opportunities in agriculture, and the availability of credit, schools, and health facilities. These data can be used in conjunction with data from the household questionnaire to analyze access to services in terms of the average distance a child has to travel to attend primary school or the proportion of households living within 20 kilometers of a hospital. The data can also be used to evaluate government programs. The price section of the questionnaire is designed to establish the local costs of food and nonfood items. One way in which these data can be used is to devise spatial price indices in order to accurately measure regional patterns of poverty (Ravallion and Bidani 1992).

Although community-level data enhance the usefulness of data collected at the household level, social scientists have only limited knowledge of how to collect community-level data. Too often in the past, the designers of multitopic surveys in both the developing and developed world have hastily assembled and administered a community and price questionnaire after the household survey was finalized, giving little

attention to the wording of the questions, to defining the entities to which questions refer, or to the selection of the informants to whom the questions are to be addressed. Frequently, no serious effort has been made to pilot and pretest the community survey. The weaknesses of the resulting community-level data have seriously limited the extent to which these survey data sets can be used to analyze the effects of government policies on individual welfare.

This chapter argues that the household and community aspects of a multitopic household survey should be regarded as two components of the same data-collection effort and should be integrated from the earliest planning stages of the survey. The policy questions that the household questionnaires in each survey are designed to answer are equally relevant to the design of the community questionnaire. The sample for the household survey influences the design of the sample for the community survey. Also, the field procedures for the two questionnaires can and should be coordinated. In some cases alternative existing sources of community-level information, such as census records or other survey data, may be available to supplement or substitute for the collection of new community data. In other cases it may be necessary to design a special facility questionnaire to be administered in addition to the community questionnaire.

This chapter discusses a number of conceptual and technical issues associated with the collection of community data in the context of a multitopic household survey such as the LSMS surveys. The first section of the chapter describes the policy issues that have been analyzed using a combination of household and community-level data. The second section discusses how to define a "community." The third section discusses how to assemble community-level data both from existing sources and from community informant interviews or facility visits. The fourth section introduces draft versions of prototype community and price questionnaires for the designers of future surveys to consider. (The prototype questionnaires are presented in Volume 3.) The fifth section explains why certain choices were made in the design of these prototype questionnaires.

### **Policy Issues Analyzed Using LSMS Community-Level Data**

Government programs often aim to provide services within a specific geographical area on the assumption

that the services benefit individuals and families living in that area. Because the community is the level to which governments often target direct interventions, it is sensible to collect data on how government programs work by measuring the extent to which the services they provide actually exist at the community level.

This section discusses the various policy issues that can be addressed with both community-level and household-level data. The availability of data at both levels affects the extent to which these analyses are feasible. The extent to which household data can be used for policy analysis depends on three factors: the outcomes and behaviors about which data are available; whether the data are cross-sectional or measure changes over time; and whether individuals can be linked to the institutions that they use. The factors that affect the extent to which community data can be used for policy analysis include the availability of data on community characteristics over time and the degree of detail of the data available on the facilities and institutions that provide services to household members.

Over half of the LSMS surveys conducted before 1997 included community and price questionnaires. Topics covered by these questionnaires typically included demographics, the economy and infrastructure, education, health, agriculture, and the prices of food and nonfood goods (Glewwe and Grosh 1998). A smaller number of surveys have collected facility-level data as well. The LSMS surveys for Ghana, Côte d'Ivoire, Pakistan, Tanzania, and Jamaica gathered particularly rich community and facility data. Several of these data sets have been used extensively to analyze how government programs have affected human resources and outcomes.

Other multitopic household surveys have also included community and facility surveys. For example, the RAND Indonesian Family Life Surveys collected data from two groups of community informants, 12 health facilities, and eight schools for the 321 communities in which household respondents were located (Frankenberg and Karoly 1995). The second Indonesian Family Life Survey revisited many of the same health providers and schools as the first, providing panel data at the facility level. Also, Demographic and Health Surveys (collected by Macro International), which have been administered in many countries, have usually included a short questionnaire about service availability.

Most existing LSMS data sets contain information on the availability of sanitation facilities, power, water supply, and public works such as road and transport networks and (in some cases) irrigation systems. The LSMS data from Vietnam have been analyzed with the aim of discovering the extent to which basic types of infrastructure are available to different population subgroups (van de Walle 1995). This analysis has revealed substantial differences between the rural populations in the north and south, particularly in terms of access to a post office (far more available in the south) and electricity (far more available in the north). However, there are only slight differences between the populations in the two regions in terms of access to piped water. Fewer than 10 percent of residents of rural areas have access to piped water, regardless of their region or their poverty status, although the poor generally have less access than the nonpoor.

The LSMS community-level data sets contain information about whether or not various types of health facilities and schools exist in the community as well as information about how far people have to travel to the nearest facility or school (and the costs they must incur to do so) if these facilities are not located within the community itself. By combining data on access to schools or facilities with household-level data on economic welfare and other basic socioeconomic characteristics, it is possible to produce descriptive statistics on households' access to health care and to educational opportunities. Such statistics have been produced using data from the LSMS survey in Ghana in 1988. The statistics show that households were an average of 0.4 kilometers from a primary school but were almost 15 kilometers from a secondary school (Lavy and others 1996).

Descriptive statistics can also be produced by region, by rural versus urban location, by economic strata, or by level of education. For Ghana these statistics show that in 1987, urban residents lived an average of 0.6 kilometers from a health facility, whereas rural residents lived almost 5 kilometers away (Lavy and others 1996). Descriptive analysis of LSMS data for Vietnam show that a high proportion of the population had access to a lower secondary school in the early 1990s and that about 80 percent of the rural population lived in communities that contained a lower secondary school (van de Walle 1995). These statistics can illustrate the extent to which governments have or have not succeeded in extending services to

the population as a whole and to different groups within the population.

LSMS household questionnaires collect data on use of services, including curative and preventive health care and contraceptive services, as well as on school enrollment. These data can be combined with community-level data on access to services to illuminate the relationship between access to services and use of services. For example, in Ghana the probability that a child age 5–12 has ever attended primary school increases by 30 percent for every one kilometer decrease in the distance he or she has to travel to reach a middle school. This finding suggests that access to higher levels of education is an important factor when parents are deciding whether or not to enroll their children in the preceding level of schooling (Lavy 1996). A similar phenomenon was observed in Vietnam (Glewwe and Jacoby 1998).

Data from the Indonesia Family Life Surveys suggest that the availability and quality of private health service providers within a community affect women's knowledge of public facilities in that community, and vice versa (Frankenberg and Beegle 1998). Data on access to facilities can be combined with data on individual characteristics such as age, sex, education, and income level to evaluate whether the relationships between access and use differ among individuals. These relationships may be of particular interest if programs are intended to benefit certain groups.

In countries where Demographic and Health Surveys have been carried out, the data sets from those surveys include data on access to and use of contraceptives. Analyses of Demographic and Health Survey data from both Tanzania and Nigeria suggest that increasing the availability of the birth-control pill in pharmacies is associated with an increased use of contraceptives (Beegle 1995 and Feyisetan and Ainsworth 1994).

Many LSMS household surveys have collected measures of human resource outcomes as well as measures of the use of services. Education-related outcomes include grade attainment and, in rare cases (such as the Ghana survey), achievement scores. Health-related outcomes include self-reported morbidity and anthropometric measures. Community-level data on access can be combined with outcome measures to explore the relationship between access to services and outcomes. For example, in an analysis of the Ghana data, Lavy (1996) found that an increase in

the distance that a child has to travel to a middle school is associated with fewer years of schooling attained.

It is also possible to use the data on labor in conjunction with community-level information on infrastructure to explore how activities in the labor force vary depending on how much and what kind of infrastructure is available to the community. Vijverberg (1995) found that in rural Vietnamese communities, important determinants of the decision to start up a nonfarm enterprise include the availability of electricity and piped water at the community level, the availability of a market that is frequently open, and the presence of a secondary school. When agricultural extension services are available in a community, this tends to encourage farming and to discourage the initiation of nonfarm enterprises.

Most LSMS community questionnaires include questions about when facilities such as schools and health clinics opened. When data are available from the community questionnaire on changes in access and retrospective data are available from the household questionnaire on specific household behavior and outcomes, it is theoretically possible to relate intracommunity or intrafamily changes in behavior or outcomes over time to changes in access to services.

For example, data on pregnancy, the use of prenatal care, and the site where the birth took place could be combined with data on the year in which a government clinic opened in the community to produce statistics on the association between the availability of the services provided by the clinic and the use of those services. If the data on pregnancy also contained data on the baby's birth weight, the analysis could additionally consider the association between the expansion of services and this important infant health outcome. The same kind of statistics could be calculated using retrospective data on contraceptive use. This author is not aware of LSMS data having been used in this way. Some of the disadvantages to using LSMS data this way include the fact that community informants may not reliably be able to remember when facilities opened. Also, analyses using this approach are not likely to be terribly informative unless access to facilities has changed in a relatively short time, because retrospective data on outcomes of interest are usually available for only about five years prior to the survey date.

Nevertheless, two analyses of Demographic and Health Survey data from Indonesia have used a version

of this approach. Frankenberg (1995) used data from the Demographic and Health Surveys and the Central Bureau of Statistics to analyze intracommunity changes in the risk of infant mortality associated with increased access to public and private facilities. This study found that there was a significant decrease in the risk of infant mortality as access to private midwifery services increased. Gertler and Molyneaux (1994) undertook a similar exercise using administrative data from the Indonesian National Family Planning Agency with respect to access to family planning services, and found that changes in family planning program inputs were responsible for 4–8 percent of the decline in fertility that occurred between 1982 and 1987.

This chapter has so far focused on combining household data on outcomes or on use of services with community informant data on access to services. However, another important factor is the quality of the services provided, and data at this level of detail can usually only be gathered effectively by designing and fielding a facility survey. For example, in the Ghana, Jamaica, and Côte d'Ivoire LSMS surveys, interviewers visited schools, health facilities, or both and collected data on the quality of services provided.

If analysts have information at this level of detail, they can analyze particular aspects of a government program. For example, Thomas, Lavy, and Strauss (1996) analyzed the Côte d'Ivoire LSMS data to find out how various dimensions of the quality of the services provided by a health facility (such as staffing patterns and drug availability) affected the height and weight for height of children. They found that as the number of doctors at facilities increased, children's heights (standardized for age) rose as well. The availability of drugs for treating common ailments was also associated with improvements in children's heights and weights for heights. By combining facility characteristics with household characteristics, the authors showed that access to immunization services significantly increases the height of poor children but has no significant effect on nonpoor children.

In some LSMS surveys, the collection of household and community-level data was organized so that data on individuals could be explicitly linked to the data on the facilities that the individuals knew about and used. This allowed analysts to examine how individual outcomes are affected by the quality of the services provided by the facility and how the quality

of those services affects an individual's willingness to travel to or pay for those services. In the context of education, Glewwe and Jacoby (1994) used data from the Ghana LSMS to show that investing in repairing classrooms is a more effective way to improve student achievement than is providing additional instructional materials or improving teacher quality.

Glewwe and others (1995) used the Jamaica LSMS survey to consider the effects of a range of physical, organizational, and pedagogical characteristics on student achievement. Overall, they found that variables that measured pedagogical processes (for example, the amount of time that students spend doing written assignments and being tested) were more important predictors of achievement than variables that measured input levels or school organization.

When panel data exist for both households and communities or facilities, it is possible to conduct analyses that relate changes in facility characteristics to changes in the behaviors and outcomes of households and individuals. These sorts of analyses have been conducted with the Indonesia Family Life Surveys. For example, between 1997 and 1998 there was a dramatic decline in the proportion of children under age 3 who had received Vitamin A in the previous six months, from 55 percent to 43 percent. Analysis of facility data from the communities of the survey respondents documented a concomitant decline in the proportion of both public and private facilities offering Vitamin A (Frankenberg and others 1998).

To date the LSMS surveys have not provided panel data on both households and facilities. The Indonesian Family Life Surveys demonstrate that with careful planning and design it is possible to do so. Perhaps in the future the LSMS and other multitopic surveys will be implemented in a manner that yields panel data at multiple levels.

This discussion has focused predominantly on analyses that combine measures of human resources from the household questionnaire with measures of access to or quality of facilities from the community questionnaire. The community questionnaire usually administered in LSMS surveys contains a number of other measures of community characteristics, such as food prices, the availability of agricultural extension services, wage rates, and transportation and sanitation infrastructure.

Analyses of these characteristics can also be relevant to policy. For example, Thomas, Lavy, and Strauss

(1996) used data from the Côte d'Ivoire LSMS survey to analyze the effects of variations in food prices on children's nutritional status. They found that in rural areas, weight for height was negatively affected by high prices for fresh fish, eggs, palm oil, and manioc, while in urban areas it was the prices of rice, sugar, and plantains that were associated with reductions in weight for height. The study concluded that the increases in food prices that typically accompany stabilization programs are likely to cause the nutritional status of older children to deteriorate.

In a similar analysis of the Ghanaian LSMS data, Lavy and others (1996) found that children's height for age and weight for height were significantly associated with the quality of water and sanitation facilities in rural communities. An analysis of infrastructure in Vietnam (van de Walle 1995) found that although the poor had less access to most types of infrastructure than did the nonpoor, infrastructure was woefully inadequate to meet the needs of either group, and any increase in infrastructure was not likely to be redistributive.

One important way in which price data can be used in combination with household-level expenditure data is to make welfare comparisons across regions. In most countries price disparities among regions are considerable. Without community data on prices, it would be impossible to compare expenditure levels among areas with different prices.

The Indonesia Family Life Surveys illustrate this point. One round of the surveys was conducted in 1997. An additional round was conducted a year later, in 1998. In the intervening 12 months, the Indonesian rupiah collapsed and prices changed dramatically. Welfare comparisons based on expenditure levels would have been meaningless without adjustments for price changes. Detailed monthly data on price changes are available from the Central Bureau of Statistics, but only for urban areas. Price data from the 1997 and 1998 community surveys were used to provide evidence that prices had risen more quickly in rural areas, and that conclusions about changes in welfare levels by sector of residence depend critically on assumptions about sector-specific inflation rates (Frankenberg, Thomas, and Beegle 1999).

Many of the papers discussed in this section have used sophisticated statistical techniques to analyze LSMS household-level and community-level data. Several problems frequently arise in policy-oriented analyses of community and facility characteristics for

which it is difficult, though often possible, to adjust statistically. First, there is the problem of omitted variables. Communities with access to infrastructure, health facilities, and schools may well have a number of other attributes, such as well-maintained transportation infrastructure, that contribute to positive human resource outcomes. If the analysis does not control for these other attributes, the effects of access to facilities will be overstated.

A second problem, particularly with analyses of facility characteristics, is that measures often tend to be highly collinear. For example, the facilities that operate during inconvenient hours are often also understaffed and understocked. Third, missing data at the individual, household, community, or facility level often mean that analysts can only use a significantly smaller and probably nonrandom subset of observations. A fourth problem is the potential endogeneity between community characteristics and individual or household-level behavior and outcomes. Governments may intentionally locate programs or resources in areas where residents have certain characteristics or households with certain characteristics may move to areas precisely because certain programs or resources are available there (Rosenzweig and Wolpin 1986, 1988). In both of these situations a straightforward regression of individual outcomes or behavior on access to programs will generate biased estimates of program impact.

### **Defining a “Community”**

One of the two most fundamental decisions in designing a community questionnaire to accompany a household survey involves identifying a basic geographical area that defines “community.” Defining the term “community” is difficult because even within countries, communities are extremely heterogeneous. A definition appropriate across the range of contexts covered by LSMS-type surveys will be too vague to be informative. What can be said is that at a conceptual level, the term “community” in the context of the surveys discussed in this book refers to a spatial unit that contains the households included in the survey sample, that has characteristics common to its residents, and that is of social, economic, or physical significance to its residents.

In LSMS surveys the definition of the community is inherently tied to the design of the household

survey. Households are grouped into sampling units (usually census enumeration districts), which are referred to as clusters. The cluster is the geographical unit in which the survey households are located. Generally, for each cluster, a “community” unit is defined that contains the households located in the survey cluster. The community data collected will be tied to that unit, and there will be a minimum of one community-level observation per unit.<sup>1</sup>

What geographical unit is the most appropriate definition of a “community” in LSMS terms? One choice is to use the cluster boundaries to circumscribe an area that would serve as the “community” for the purposes of administering a community questionnaire. In most survey contexts, clusters are small, contiguous units that do not overlap. However, the cluster boundaries are typically defined by the central statistical agency of the country and are often too small or too arbitrarily determined to be socially, economically, or physically significant to the people living in those areas.<sup>2</sup> This lack of significance is the main disadvantage to using cluster boundaries to define the community.

A second possibility is to define the community solely in terms of its size—for example, by the area contained by a radius extending five kilometers from the center of the cluster. The apparent simplicity of this method and its uniformity across clusters are the main advantage of this approach. The biggest disadvantages is that the selected radius is bound to be arbitrary and may not correspond to the informants’ notion of the community to which they feel they belong. In addition, this definition is difficult to put into practice if there is no clear center to the cluster. If the community is organized along a road or stream or if it borders a body of water, a circle is not an appropriate shape. Moreover, if informants do not have a good sense of distance and direction, they may not give clear and accurate answers to questions about the area.

The third possibility is to define the “community” in terms of the administrative unit or units of government under the jurisdiction of which the survey households fall. In most countries this can consist of several different levels of government.<sup>3</sup> In previous LSMS surveys, clusters have tended to be small enough to be contained within a low level of government. Therefore, the community questionnaires in most of these previous surveys have defined the “community” as the most appropriate low-level administrative unit, for the following reasons:

- An administrative unit is a well-defined geographical area to which informants can easily relate.
- Measuring access to government programs is often a primary goal of the community questionnaire in LSMS-type surveys, and the benefits of government programs are often allocated by administrative unit.
- Administrative units are usually run by people whose responsibilities imply that they are knowledgeable about many of the topics dealt with in a community questionnaire.
- Some data broken down by administrative unit may already exist at the central government level.

The most sophisticated approach would be to define the “community” in accordance with both the particular phenomena of interest to the survey designers and the characteristics of the household survey cluster. It is important to bear in mind that it may not be immediately apparent which definition is most appropriate in a given situation. Also, one definition of community will not always be equally applicable to all clusters or to all of the characteristics of interest. Therefore, it is not essential to adhere to one definition of community in all parts of the community questionnaire or in all contexts. For example, survey designers might choose to define the “community” using village boundaries in rural areas and postal codes in urban areas. Some variables, such as access to piped water systems, may best be captured at a very local level, while other variables, such as distance to the nearest hospital, may be better suited to a wider definition of community.

The best way to measure the effects of community-level characteristics on the outcomes and behaviors of survey respondents varies depending on the mix of characteristics exhibited by the community. This variation has consequences for the way in which the questions about each characteristic are worded in the questionnaire. For example, there are more pharmacies than hospitals in most developing countries, and pharmacies typically serve clients from a far smaller catchment area than do hospitals. Consequently, when information on the availability of pharmacies is being sought, the most appropriate question to ask community informants may be whether one or more pharmacies is located within five or ten kilometers from the community center. In the case of hospitals, it may be more relevant to ask informants how far they live from the nearest hospital.<sup>4</sup>

Community-level characteristics were measured in various different ways in previous LSMS surveys

depending on whether the cluster studied was located in an urban or a rural area. In several surveys, almost no community-level data were collected in urban areas,<sup>5</sup> which seriously limited the extent to which those data could be used for policy analysis.

One argument against collecting community data in urban areas is that urban residents have access to such a wide range of resources that the location of a household is an insignificant determinant of the opportunities and constraints faced by its members. Therefore, it is argued, collecting information on the surrounding area is not very informative. This view is short-sighted. There is no conclusive evidence to suggest that urban residents are completely unaffected by measurable aspects of their environment. Also, this argument ignores the fact that, in reality, the urban/rural divide is a continuum rather than a clear-cut distinction. Thus, rather than deciding in advance that community characteristics are irrelevant in urban areas, survey designers should structure the community questionnaire in such a way as to avoid having respondents answer sections that have little relevance for the community in which they live.

A second reason why urban areas may have been excluded from community modules in previous LSMS surveys is that it is particularly difficult to define what is meant by a “community” in an urban context. This is undoubtedly true, although the structure of local government may be a good guide in some countries while elementary school catchment areas or postal zones (with accompanying maps) might be appropriate in others.

There may be some situations in which the decision not to collect community-level data in urban areas is sound, but this decision should be made on a survey-by-survey basis and only after careful consideration. When making this decision, survey designers should bear in mind the policy issues that the survey is designed to illuminate, the administrative structure of urban areas, and the availability of other sources of information about urban areas.

The definition of a “community” can also be unclear in areas where the population is extremely sparse and where the survey households are spread out rather than “clustered.” If there is no obvious “community center” that is a focal point for community life (perhaps a place where services are provided or goods can be purchased), each household is effectively its own community, which may mean that there is no

point in collecting community data. This scenario may seem extreme, but in practice there have been some cases in which settlements were so scattered that there was no point in including questions about distances in the community questionnaire.

### **Assembling Community-Level Data**

Having established the definition of the communities to be studied, survey designers should then investigate the various possible ways of gathering the pertinent community-level information. Broadly, there are three ways to assemble community data:

- Using any existing (or “secondary”) data from administrative archives or previous studies.
- Conducting community informant interviews.
- Visiting facilities, service points, or markets.

“Secondary” data have already been collected and assembled, whereas the other two categories require the survey team to gather new information. None of these data sources alone is likely to provide all the community-level information needed, but, when combined, they can paint a detailed picture of the community. The designers of a multitopic household survey should thoroughly explore the feasibility of each of these three methods of assembling community data at the earliest stage of planning.

It is also possible to use data gathered in the household questionnaire of the survey to generate some community-level characteristics. For example, by calculating the average wage of the respondents to the household questionnaire, it is possible to construct a measure of community-level wage rates. Another example would be to construct a measure of the community’s access to piped water by calculating the proportion of survey households that have piped water from responses to the household questionnaire.

The usefulness of these community-level aggregated measures depends on the size of the geographical area to be characterized, the number of respondents per community, and the degree of heterogeneity across potential sampling units within that area. If the aim is to characterize only the area encompassed by the sampling unit from which survey households have been selected, aggregating information from a set of randomly selected households within that area will produce an unbiased estimate of the phenomenon of interest. However, if the number of respondents per community is small, any community-level estimates

obtained from aggregating household responses are likely to be biased. If the aim is to characterize an area larger than the sampling unit, aggregating data from the household questionnaire will not produce an unbiased estimate of the phenomenon if there is substantial cross-cluster heterogeneity.

This issue can be made more concrete with two examples. Suppose the phenomenon of interest is the attractiveness of an area to migrants. Wage rates are one potential measure of attractiveness. If the survey households are poor and are located in a relatively poor neighborhood, the low community-level wage rate obtained from aggregating the household responses will be an inaccurate measure of the wage rates and thus the attractiveness of the surrounding area. On the other hand, suppose the phenomenon of interest is the degree of exposure to household waste in the immediate vicinity of the survey households. In this case, aggregating from the household data may provide an accurate measure of exposure within the small area the analyst wishes to characterize.

### *Sources of Secondary Data*

Because serious efforts have been made in many developing countries to collect or assemble data on a wide range of topics, there is considerable potential for supplementing LSMS surveys with secondary data.

Survey designers may find that secondary data exist on, for example, weather patterns, food prices, housing characteristics, wage rates and occupations, the quality of schools and health facilities, and the distribution of family planning methods. In Vietnam, for example, it was not necessary to collect price data from urban areas in the LSMS survey because the statistical agency was able to provide these data. It is well worth exploring whether such data exist and evaluating whether an existing source can meet the survey’s community-level data needs. When secondary data are the sole source of community-level data, the choice of the geographical unit to which those data apply will depend primarily on what data are available.

The survey team should take the following basic steps: identifying relevant secondary data and assessing their quality; obtaining permission to use the data from the organization that has assembled them; and ascertaining whether it is possible to merge the secondary data with the data from the household questionnaire of the LSMS-type survey based on a geographical link between the survey cluster and the unit

to which the secondary data refer. If data are available at the level of individual institutions (for example, for each school), survey designers need to decide which schools are relevant to individuals living in the survey's sample households.

**IDENTIFYING RELEVANT SECONDARY DATA.** There are several questions that survey designers should ask themselves to determine whether the available secondary data will be a useful supplement to an LSMS household survey. The first question is whether the data cover the entire country or only selected geographical areas. If the data are limited geographically, then the survey designers must assess how close the LSMS clusters are to the sites covered by the secondary data set. In some cases, the secondary data may be so valuable that it is worth trying to select the LSMS clusters so that they can be linked to the secondary data. For example, in a survey focusing on agriculture, the value placed on having weather data may be so high that it may make sense to select clusters for the LSMS survey that are located near a weather station.

Another question to bear in mind is the age of the available data; some—such as diennial censuses—may be so old that they are of little use. Additionally, survey designers should ascertain whether the data are gathered regularly (as in a monthly price survey) so that a time series of data is available.

Even when the geographical coverage and collection date suggest that the secondary data will be useful, the quality of the information should be investigated further. Many secondary data sets sound better than they turn out to be in practice. For example, the fact that every health center is required to turn in a monthly report of activities is no guarantee that each center actually does so. There may be weather stations throughout the country, but their equipment may not function. If the data are incomplete, the designers should assess the importance of the missing data and whether the data are recoverable.

Secondary data often do not contain sufficient information from which to construct variables that enable analysis of the relevant policy questions in a country. For example, if a high priority in the health sector is training personnel and assigning them to remote areas, it may be useful to evaluate whether the use of facilities has increased or health status has improved in those areas since the number of staff increased. However, the secondary data might only

include information on, for example, the bed capacity of health facilities. This kind of issue poses more of a problem for policy researchers than for academics, since academic researchers have more scope for shaping their questions to fit the available data.

A related concern is whether not only the key program variables but also “control” variables, such as basic measures of infrastructure and socioeconomic development, are available in the data. Because there are often many correlations among the various development indicators, failing to control for other types of infrastructure in the community may result in misleading results because of omitted variable bias. (See Chapter 26 on econometrics.)

Another key issue concerning secondary data is whether they contain too much measurement error to be worth the difficulties involved in using them. Measurement error, which can be random or systematic, may arise in a number of ways. First, the data may have been collected at too high a level of aggregation to capture the aspects of the environment that affect individuals' behavior and outcomes. For example, it may be possible to merge administrative data on regional governments' expenditures on family planning into a household survey data set, but those expenditures may be so weakly correlated with the aspects of family planning services that affect an individual's contraceptive use that they have no explanatory power in a model predicting contraceptive use.

Second, secondary data may not reflect the conditions that actually prevail in a given community or facility. For example, using data collected through visits to facilities, Thomas, Lavy, and Strauss (1996) showed that the estimated impact of health service infrastructure on child anthropometry is much smaller when the infrastructure is measured in terms of what is supposed to be available at a health facility rather than in terms of what is actually available. Quality issues are particularly troubling when the differences between reality and what is reported are systematic (nonrandom measurement error) because administrators have an incentive to report conditions as being particularly good or bad.<sup>6</sup>

Finally, survey designers should consider whether the available secondary data cover only public sector facilities and not the facilities run by the private sector, which is common. The importance of such an omission depends on the size and geographical distribution of the private sector facilities and on whether

the services they provide compete with public facilities for clients. For example, in some countries private elementary schools may be so rare and prohibitively expensive that omitting them from an analysis of the enrollment decisions of low-income families poses few problems. On the other hand, private health care facilities may be an important source of health care for the populations of many developing countries.

**GETTING PERMISSION TO USE SECONDARY DATA.** Sometimes acquiring secondary data can be time-consuming and expensive. Agencies may be hesitant to release data. However, almost all previous LSMS surveys have been implemented by the national statistical agencies of the countries studied, which has usually made it easy for the survey team to at least obtain information on other sources of official data.

**MATCHING THE SECONDARY DATA TO HOUSEHOLD SURVEY DATA.** Combining secondary data with household data is theoretically straightforward, but in practice several problems can arise along the way. A different geographical coding scheme may have been used in the secondary data set than the one used in the LSMS-type survey. Incompatible coding schemes are more likely to be a problem if the supplementary data are obtained from an organization other than the one conducting the household survey.<sup>7</sup> Ideally, both data sets should record the names of the administrative areas (or facilities if that is the level at which the matching occurs) about which the data were collected as well as a common set of codes. Even if a common set of codes is available, the matching should be verified by name.

Secondary data certainly have considerable potential as a source of information on the communities and the larger administrative units that surround the household survey clusters. However, because of the numerous problems that can arise in tapping that potential, survey designers should investigate fully the extent to which existing sources of data can meet the survey's policy research priorities. It would be foolhardy to assume at the outset that the existence of secondary data precludes the need to collect community-level data in the survey.

#### *Conducting Community Informant Interviews*

A second way to obtain community-level data is to interview one or more residents of the community

who are knowledgeable about the environment that is common to all the households located in the community. This method is easy to implement and has the additional advantage of being cheap.

**SINGLE-INFORMANT INTERVIEWS.** In its simplest manifestation, this approach involves the survey team conducting an interview of no more than a couple of hours with the community informant after administering the household questionnaire to the sample households in the area. In many surveys the community informant is likely to be the community leader with whom the survey team would have to meet anyway, as a courtesy or to obtain his or her permission to conduct the household interviews.

When the survey designers decide to use informant interviews to collect community data, the community should be defined according to the type of information needed and the feasibility of identifying and interviewing informants who are well informed about that community. Both the geographical unit to which the community data pertain and the sources of data should be chosen so as to match the priority topics of the overall survey. A survey with a special emphasis on infant and child health will require a different set of community-level data than a survey that emphasizes secondary school attendance or one that focuses on poverty.

Although this method has its problems, it is a good way to obtain a "core" set of comparable information tailored to the purposes of the household survey about the communities in which the clusters are located. It is useful to have this information because many analysts will want to evaluate the effect of one community characteristic while controlling for other characteristics with which it is likely to be correlated. The biggest determinant of the success of this approach is the knowledge level of the informant relative to the questions being asked. It is counterproductive (though tempting) to ask questions at a level of detail about which the informant is likely to be ignorant. Thus it makes sense to design informant selection protocols that anticipate what types of informants are likely to be knowledgeable about which topics.

**GROUP INTERVIEWS.** The practice followed in most previous LSMS surveys has been to assemble a group of informants consisting of, for example, village chiefs,

teachers, government officials, and health care workers, and to administer one community questionnaire (composed of various modules) to all these informants at the same time. If retrospective data are needed, it is useful to include at least a couple of residents who have lived in the community for a number of years. Interviewing several people instead of just one is a sound approach because each community member is likely to be well informed about a different topic. If the members of the group represent different areas of expertise, the most knowledgeable member can take the lead in answering the questions about his or her topic.

If this approach is used, it should be formalized in the community survey protocols. Interviewers should be given explicit instructions about how to identify informants and assemble a group. They should be instructed about what types of informants must be included (for example, a teacher or someone active in agricultural extension activities). If there is an obvious community leader or respected elder, this person may be enlisted to organize a group of informants. The questionnaire should provide spaces for recording the basic characteristics of each informant, including his or her name, sex, age, education, position in the community, length of tenure in his or her current position, and length of time he or she has lived in the community. At the end of each module, the interviewer should record the identities of the informants who have participated in the discussion.

One problem with group interviews is that they may be hard to control. If the group of knowledgeable informants is composed of people with such different backgrounds or interests that it is unlikely they will reach any consensus, it may be better to conduct separate interviews with individual informants or to interview several smaller groups of informants. For example, a woman who volunteers to distribute contraceptives from her home may be a good source of information about family planning services in the community. However, if she is likely to defer to male community leaders in a large group interview rather than speak up, she should be interviewed as part of a separate group, possibly with other knowledgeable women in the community.

Groups can be organized according to, for example, employment status (separating government workers from nongovernment workers), leadership status (interviewing elected or appointed local leaders or

community elders), gender, or the length of time people have lived in the community. One advantage of interviewing more than one group per community is that it is then possible for the survey team to compare the different interview reports to find out in which clusters and on what topics the groups disagreed. These comparisons may suggest which community-level variables are particularly subject to measurement error (as indicated by disagreement among the different groups) and whether certain clusters (such as those in urban areas) are more prone to measurement error than others.

Conducting more than one group interview allows for a heterogeneity in responses that might not emerge in one group interview with diverse members. Another way to encourage heterogeneity is to allow multiple answers to some questions in the questionnaire. Often, community surveys ask about the primary source of *X* (for example, drinking water), the main type of *Y* (for example, road surface), or whether a *Z* exists (for example, a pharmacy). If there are many different sources of drinking water, types of road surfaces, or numbers of pharmacies, it may be preferable to ask informants both to specify all the relevant options and to identify and provide more detail about the primary option.

**MEASURING THE AVAILABILITY OF SERVICES.** One of the principal objectives of a community survey is to establish the number and quality of services—such as schools, health facilities, contraceptive outlets, banks, and markets—to which community residents have access. The content of this part of the questionnaire depends on the extent to which survey designers want to learn about residents' options and on how much detail they want to collect about the services. These factors also determine whether it is necessary to collect data at the facility level.

First, a question arises: what constitutes "available?" In this chapter, "available" facilities are ones that most people interviewed by the household survey are aware of, if only vaguely, and would at least consider using. The number of facilities that meet this definition is likely to vary among communities and among types of facilities, as is the geographical size of the area that contains them. Other possible ways to define "available" include facilities that meet some geographical criterion, such as all facilities within a certain radius of the center of the cluster or all facilities within the administrative unit that contains the cluster.

This section describes three ways of measuring service availability:

- Asking community informants whether a type of service or facility is available.
- Asking community informants to identify the main facilities to which households in the community have access.
- Combining responses from the household questionnaire with information provided by community informants.

Questions about the availability of services may aim to establish whether a particular service is located within the community's boundaries. Such questions may also aim to establish how far the service facility is from the geographical center of the community or the time it would take to travel that distance and how much it would cost to do so. (See Section 3 of the draft community questionnaire in Volume 3.) In some previous surveys, community informants have been asked questions about both topics.

Generally, it is easier to ask about the availability of a service within a given boundary than to try to obtain estimates of distance. However, the significance of whether a service exists within the community boundaries depends on how the boundaries are defined. If the boundaries refer to a very small neighborhood (as may be the case for census enumeration areas), the existence of a service within those boundaries is less significant than if the boundaries refer to a larger area. On the other hand, if the boundaries refer to a very large area, it will not be possible to gauge the true availability of the service to small communities within that area. For each type of service facility (for example, a health clinic versus a hospital), survey designers should carefully consider which existing administrative unit to choose in defining community boundaries to ensure that the resulting data are meaningful.

Another possibility is to ask community informants about the distance either to the nearest facility or to a finite number of facilities of specific types (see Section 6 of the draft community questionnaire in Volume 3). In questions on distance, the term "community center" is often used to denote a reference point from which distance is calculated. The more specific this reference point, the better. If communities typically have a gathering point such as a town hall, a place of worship, or a market, this location can serve as the reference point. However, if households in the community are widely dispersed but it is only a short

distance between the community center and the service facility, this distance obviously does not reflect the access that the members of these scattered households have to this facility. In some countries, it may be possible for survey designers to use Global Positioning System technology to make objective measurements of distance by visiting the facilities. (This is described further below in the section on visiting facilities to collect data.)

If survey designers want to characterize the extent of choice community members have among different facilities providing the same service—for example, several different health clinics—it will be necessary to obtain information about multiple facilities or service points. (See Sections 9 and 10 of the draft community questionnaire.) Once designers have defined "availability," they need to identify all of the facilities in the community that meet that definition. One way to do this is to explain the criterion to one or more community informants and ask them to identify the facilities that qualify according to this criterion. Each facility that the informant identifies can be listed on a service availability roster.

Another possibility is to ask household respondents to identify facilities during their household questionnaire interview. Their responses can be compiled into a cumulative list for the cluster as a whole. This method requires that the household respondents identify each facility by its name and address and that the interviewer compile the responses into one list. Both of these activities are time-consuming. The cumulative list of facilities for the cluster can then be verified with one or more community informants to check whether any obvious facilities are missing from the list and to add information about the distance and the travel times and prices for each facility on the list. If this method is used, the list cannot be compiled until all of the interviews for the household questionnaire are finished.

Which method is preferable depends on the goals of the survey and on whether data are to be collected from the facilities themselves. If the community survey involves interviewers visiting facilities, it is recommended that households be asked to identify facilities precisely in the household questionnaire and that these responses be compiled into a list for the cluster. (The advantages of this approach are described below.) This list, supplemented by information from one or more community informants, also provides

data on service availability. If facilities are not to be visited, survey designers must decide how thoroughly they wish to investigate the range of facilities available to households. Community informants will probably be able to identify the main services that are easily accessible to members of that community. They are not likely to provide a complete list or much detailed information about those facilities because they simply will not know everything about all facilities. Analyses of the Indonesia Family Life Survey data have revealed that community informants do a better job of characterizing public facilities than private facilities (Frankenberg 1998). If detailed information about the facilities is desired, interviews should be conducted at the facilities.

#### *Visiting Facilities to Gather Data*

A third way to obtain community-level data is by visiting facilities and administering questionnaires to staff. The term “facility” is used broadly and includes markets and sales outlets (for gathering information on prices) as well as schools, health facilities, banks, other sources of credit, and employers.

The main advantage to visiting facilities is that it becomes possible to obtain far more detailed and generally more accurate information about the prices and content of the services that these facilities offer than can be obtained just by asking community informants. By making direct observations of these facilities, it is also possible to gather information about private as well as public facilities.<sup>8</sup> Potential topics for facility questionnaires are covered in the respective sectoral chapters of this book.

**SELECTING FACILITIES FOR INTERVIEWS.** One of the first issues to be considered when designing a facility survey is how to choose which facilities will be visited. This issue is a logical extension of the discussions regarding how to define “community” and how to measure service availability. In an LSMS survey, facilities of interest are facilities that are “available” to the households in the LSMS sample. Generally, survey designers aim for a good deal of overlap between the facilities interviewed and the facilities that LSMS household respondents know about and use.

From a theoretical standpoint, the way to select facilities that parallels the typical procedure for selecting households is to define a geographical area of interest, to list all the facilities in that area, and then to

select a sample from that list. However, this strategy is rarely feasible, for the following reasons:

- It is difficult to know a priori what constitutes an area of appropriate size.
- The listing process is time-consuming, expensive, and probably impractical to undertake in an area large enough to cover the distances household members are willing to travel for services.
- The facilities chosen for the sample may not correspond to the facilities that LSMS households know about and use.

The strategy recommended here is to compile a list of facilities in the household questionnaire as described above. If the number of facilities on the cumulative list is small enough or if the survey budget is large enough, all the facilities on the list can be visited. Otherwise, a sample of facilities may be drawn from the list. The sample of facilities can be selected either randomly or according to some other criterion—for example, in proportion to the number of times the facilities are mentioned by household respondents.

This was the procedure used to draw a sample of facilities in the Indonesian Family Life Survey. Facilities were ranked according to the frequency with which they were mentioned by survey respondents. The most popular facilities were visited, and additional facilities were randomly selected until a predetermined quota was reached. This method guaranteed that all facilities had a nonzero probability of being selected, while increasing the chance of substantial overlap between the facilities interviewed and the facilities of relevance to household residents (Frankenberg and Karoly 1995). Information on the number of facilities listed and how often each was mentioned by household respondents was used to generate the sampling weights for the selected facilities (McCaffrey, personal communication, 1995).

From a theoretical standpoint, using a cumulative list generated from the responses of household members is attractive. The main drawbacks are that the list cannot be constructed until household interviewing is finished and that compiling the list may be time-consuming, depending on the number of household respondents and the number of different kinds of facilities the respondents list.

Other methods that have been used to select a sample of facilities are to select all facilities within a certain radius of the cluster center or all facilities with-

in the administrative unit that contains the cluster. Yet another alternative is to choose the facility located closest to a cluster. This strategy will almost certainly generate a biased picture of available services in any community in which residents have access to multiple facilities. All of these strategies are flawed from a scientific perspective and present analysts with a number of difficulties.

**USING GLOBAL POSITIONING SYSTEM TECHNOLOGY TO MEASURE ACCESS TO SERVICES.** Visiting a sample of facilities makes it possible to use Global Positioning System technology to measure the distance between a facility and the center of the community or cluster. The Global Positioning System is a navigation system that can be used to determine a position on the earth in relation to a set of orbiting satellites. The Global Positioning System can determine the latitude, longitude, and altitude of the facility, the coordinates of which can be labeled and stored for lengthy periods in the system's memory until a computer is available into which they can be downloaded directly. Global Positioning System indicators are now available quite cheaply. Most are easy to use. The operator finds a location where the sky (and as much of the horizon as possible) is visible, turns on the indicator, and holds it up until three or more satellites have been "acquired." At this point, the indicator gives a readout and the operator can mark and label the spot. There are several potential sources of error associated with the system, such as the blockage or reflection of satellite reception by buildings or other obstructions.

Global Positioning System data can be used in several different ways in the context of a community-facility survey. The most basic application involves taking readings of the latitude and longitude of each facility visited. Coordinates from the facilities, in combination with readings on the latitude and longitude of the cluster center, the community center, or (preferably) both, can be used to calculate objective measures of distance to each facility.<sup>9</sup> However, these estimates measure the most direct route, ignoring the fact that people typically travel on roads or paths rather than overland. Thus another more sophisticated way to use Global Positioning System technology to measure the distance from a central point to a facility is to take a sequence of readings along the route and to add up the lengths of the various segments of the route traveled to the destination. This method is more time-

consuming but it gives a more accurate estimate of the distance between a facility and the community center than the straight-line measurement.

A third way in which Global Positioning System data can be used is in conjunction with digitized maps of the areas in which the survey is being conducted. If digitized maps are available, the facility and cluster coordinates can be added to the map to illustrate the locations of the services available to households in each cluster. The more features, such as roads, bus stops, and markets, that are geocoded into the map, the more thoroughly it is possible to analyze any geographical factors that may prevent people from using the facilities located in their community. Survey designers should explore whether the country's mapping agency produces digitized maps and at what level of detail. This method of using the Global Positioning System was used by Entwistle and others (1997) to map contraceptive facilities in Nang Rong, Thailand. Because a detailed map of the area was available, researchers were able to calculate the time it took to travel to these family planning facilities using various different routes. The analysis revealed that the composition of the road (for example, asphalt or dirt) had an independent effect on contraceptive behavior.

#### *Collecting Price Data*

It is standard practice for LSMS surveys to collect price data by sending survey teams to visit markets. Prices are an important element of community-level data, but collecting price data is complicated, and many analysts of community-level price data have expressed dissatisfaction with the information they have been given. From the analyst's perspective, one major problem is that price data are often not sufficiently comparable among communities, nor are they sufficiently precise within communities. Three general problems account for this lack of comparability. First, price data typically contain large numbers of missing values, so that a complete set of price information is rarely available for any one community. Second, there are often differences among communities in the quality of the item for which the price is reported. Third, quantity is often measured imprecisely, making it impossible to calculate a unit price.

In thinking about how to avoid these problems, it is useful to start by considering how price data will be used. There are two major uses for price data. The first is to develop price indices, which are used to ensure

that expenditures in different regions can be compared without being affected by any price differences between those regions. The second major use for community-level prices is in models predicting behavior such as school enrollment and outcomes such as health status, because price levels are considered outside the control of households or individuals.

These uses of price data have three implications for the design of the price module. First, the items in the community price module—particularly the food items—should complement the items in the household consumption module. Second, the price module should gather price data on items—which can include food items, nonfood items, and services—that may affect the behavior and outcomes of interest in the household survey (for example, purchases of items such as aspirin, antibiotics, and condoms). Third, it is important to choose sources of information that accurately represent the prices community members encounter.

It is possible to design the price questionnaire so as to avoid some problems of missing values, variations in quality, and meaningless or imprecise measures of quantity. However, an equally important way to avoid these problems is through training and supervision of the interviewers. The filled-in questionnaires should be reviewed by field supervisors; if they are incomplete or imprecise, the supervisors should ensure that the interviewers clarify the information in question.

The content and design of the price module should reflect in large part the conditions of the country in which it is being used, including the usual diet of the population and the ways in which goods are distributed, marketed, and sold.

**SOURCES OF INFORMATION ON PRICES.** In most previous LSMS surveys, interviewers have collected price data by visiting markets and vendors and asking the price of particular goods. For items normally bought in bulk, interviewers would ask the price for a particular weight, with the item being weighed on scales and sometimes in containers carried by the interviewer. The interviewer would repeat this process at several sales outlets with the goal of obtaining three price measurements per item. Because not every outlet sells every item on the questionnaire, the interviewer might have visited far more than three outlets to obtain three price measures for every item.

In some instances survey designers may wish to collect specific information about each outlet that the

interviewer visits. These data, such as the name and address of the outlet, could be collected on a separate “outlet roster” form similar to the roster forms used for schools and health services. Another possible way to collect prices would be to ask community informants or a sub-sample of household informants about prices. Given how little is known about how to collect data on community-level prices and how many problems there have been in past LSMS surveys, it is recommended that both methods be used.

The outlets selected for interviewer visits should be places that a typical household respondent would use. The range of possible methods for selecting price outlets is parallel to the range of methods for selecting facilities. One method would be to record in the consumption module of the household questionnaire the market or store where the respondent usually shops for food. The interviewer could then visit the two or three outlets mentioned most often. Otherwise, the outlets could be selected on the basis of proximity to the cluster center or to some other center of activity, such as the community center or a major employer.

It is best to administer the price questionnaire to informants who frequently make purchases and, if possible, to avoid relying on informants who are considerably poorer or wealthier than the household respondents seem to be. A community-level interview conducted with a group of female informants may provide a good opportunity to ask questions about prices and other community-level topics.

**COLLECTING DATA ABOUT ITEMS OF A SPECIFIC QUALITY AND QUANTITY.** When interviewers collect data on the price of any given commodity, they must record the price for a particular quantity of the commodity using a well-known unit of measurement. This is straightforward for commodities that are routinely sold in known quantities such as liters or grams. Survey designers should specify to interviewers in advance their preferred quantities and units of measurement. If the price data are to be collected at markets, interviewers can bring scales and containers to standardize the quantities of the various commodities.

Other commodities tend to be sold in tins, cans, bunches, bundles, sachets, or packets. In these cases it is critical to record the quantity in standardized units such as liters or grams. Some countries or regions use uncommon measurement systems, but this problem can be solved by having the interviewer use his or her

own set of scales and measuring containers to quantify the commodity. The interviewers' training, fieldwork, and field supervision should all stress the importance of obtaining price data in known quantities.

Standardizing the quality of the commodities for which prices are collected is also important to ensure that analysts can later compare like with like. However, standardizing quality is much more difficult than standardizing quantity. Appendix 13.1 suggests ways for interviewers to specify the quality of items.

With respect to both quantity and quality, it is important that any relevant country-specific information be reported on the price questionnaire. If the price of rice varies depending on whether it is an imported, modern or local/traditional variety, the variety should be specified in the questionnaire. The question then arises: if a particular variety or brand is not available, should the price of another variety or brand be collected instead? The questionnaire proposed in this chapter provides a space in which survey designers can inform the interviewer about which quantity and variety they prefer. If the interviewer is unable to obtain information about that quantity or brand, he or she should record information about another quantity and variety while specifying on the questionnaire sheet the alternative quantity or brand name.

**CONTENT OF THE PRICE QUESTIONNAIRES.** There are three broad categories for which it is desirable to collect price data: food, nonfood goods, and services. The items in the household expenditure lists provide some more detailed guidance, as do country-specific lists of goods whose prices are used in calculating the consumer price index.

The food price list should cover staple starches, fish, meat, poultry, beans, lentils, fruits, vegetables, salt, sugar, milk, and cooking oil. Spices, alcoholic beverages, and tobacco are other items that may be important in some country circumstances. The numbers of items per category and the specific items in each category will vary by country, but the list should cover enough items to account for 80–90 percent of households' expenditures on food. Survey designers should ensure that the list takes into account any regional variations in dietary patterns. If chicken is the main commodity in one region while fish is the main commodity in another, the prices of both goods should be collected in both regions.

Given that the primary goal of the LSMS is to monitor poverty trends, the list should probably cover the items in the food bundle that the government uses to calculate poverty lines.

In many situations the quality of food and the way it is processed vary considerably in ways that are reflected by variations in price. When interviewers collect price data on meat, fish, and poultry, they should specify whether the recorded price applies to meat with or without bones. Also, because different qualities of staple starches, particularly rice, are often available, it is important to specify the kind of starch to which the price refers. It may be possible for survey designers to specify a particular variety that is available throughout the region or to ask about starches of different qualities. If food prices are known to vary by season, it may be useful to try to collect price data for a specific season and to establish whether an item is available year-round.

The list of prices for nonfood items can be long or short. General categories include prices for public transport to key locations, the cost of fuel, prices for basic clothing such as shirts, skirts, and trousers, housing prices (typically the rental and purchase prices for an average house in the community), prices of household goods (such as soap, cooking fuel, or firewood), basic medicines, education inputs such as school uniforms and basic supplies, and agricultural inputs such as fertilizer and insecticide.

If price data are collected from community informants rather than from visits to markets or shops, nonfood prices should be included in other parts of the community questionnaire. For example, questions about the price of fuel might be included in the section on transportation, while questions about the prices of agricultural inputs could go in the section on agriculture. This principle is important given that different informants contribute to different sections in accordance with their specific expertise. There is no guarantee that the informants who know the most about food prices also know the prices of insecticide and kerosene.

It is also worthwhile to collect price information on such services as schooling (registration fees), health care, small loans (interest rates), and sewage hookups, garbage pickup, and electricity, if these vary among communities. Since services are not sold at the same outlets as food and nonfood goods, data on the prices of services may need to come from the community

informants. For this reason, several sections of the community survey are likely to contain questions about prices distinct from the questions in the price questionnaire.

### *Summary*

This section has discussed three ways of obtaining community-level data: from secondary data, by interviewing knowledgeable community informants, and by visiting local facilities and markets. Survey designers should evaluate all of these potential methods during the design stage of the household survey. If there is an abundance of high-quality administrative or survey data that designers can use, this will obviate the need to include a full community survey of informants, markets, and facilities in the overall household survey. However, secondary data are often less useful than they may first appear, so survey designers should examine their content, quality, level of disaggregation, and comparability with the household survey data very carefully before concluding that they need not field a full community survey. One important question that should not be overlooked is whether the secondary data include any data about private-sector facilities, as these are an important element in calculating the effects of government programs on individuals' human capital investments and outcomes. Fielding a facility survey as part of the overall household survey provides an opportunity to collect data on private as well as public providers of services.

The second way to obtain community-level data is to interview one or more groups of knowledgeable informants. This is a cheap and straightforward way of obtaining information on a basic set of community-level variables that are comparable across clusters and that refer to the time period during which household-level data are collected. Therefore, it is recommended that this practice continue in future multitopic household surveys, although survey designers should make sure in advance that community informants (either individually or in groups) will be able to provide the necessary information. It is also crucial that the community questionnaires be pretested and revised before they are used in a full-scale survey exercise. In the past, this important step has often been omitted.

If one of the aims of the household survey is to collect detailed community-level data on services or facilities, a third method can be used: sending survey interviewers to visit local facilities and markets to

gather relevant information. In most previous LSMS surveys, community-level data on prices have been gathered by visiting sales outlets. This practice should continue. Other relevant service points include public and private health providers (including, where appropriate, traditional practitioners), elementary, junior, and senior high schools, and banks and cooperatives. Survey designers must first choose and test a method for selecting the facilities to be visited. Also, they should design different questionnaires for the different types of facilities, striking a balance between using content that is sensible for a particular type of facility and obtaining comparable data among facility types.

The main disadvantage of gathering information by visiting facilities is the high cost involved. Fielding an in-depth facility survey is far more expensive than interviewing informant groups and visiting a few markets. Thus it probably only makes sense to carry out a facility survey if the main aim of the household survey is to understand the determinants of investments in human capital and subsequent outcomes. If this is the case, the household survey should contain detailed measures of behavior and outcomes related to human capital, while the community survey should include visits to facilities that are thought to influence that behavior and those outcomes.

The easiest way to think about the costs associated with collecting facility-level data is in terms of the interviewer's time; such time calculations are analogous to those for the household survey. If data are needed from 10 facilities per cluster, 10 facilities must be selected and interviewed. This will take at least 10 times as long as conducting one interview with a group of community informants, but will take less time than conducting a lengthy interview with each member of 15 or 20 households. Thus an extensive community-facility survey might cost up to one-third of the price of a household survey, while a rudimentary interview with community informants will cost only about 5 percent of the price of a household survey.

### **Draft Community Questionnaire**

This section introduces a prototype LSMS-type community questionnaire (presented in Volume 3) that includes a specific set of questions on prices. The questionnaire assumes that a core set of community-level data will be collected from informants within the community, but that there may also be scope for

obtaining data from administrative sources or from visits to service points. The questionnaire proposed here covers most of the sectors on which there are modules in the household questionnaire. The sectoral components were developed with input from the authors of the sectoral chapters of this book.

Individual sections can be expanded or contracted to reflect the particular policy focus of the household questionnaire. Within the questionnaire, items that may not be essential have been marked with an asterisk to indicate that survey designers may wish to drop them in order to shorten the questionnaire. If all of these items are dropped, the questionnaire shrinks from about 35 pages to about 20 pages. Of course, the specific context and purpose of the survey should be considered when deciding which questions to drop and which to retain. For example, if the household survey focuses on education and health, it may not be necessary to include lengthy sections on credit or on sanitation infrastructure. The questionnaire should be pared down during the design and pretesting stages if it is not possible to identify respondents who are likely to know the answers to all of these questions.

While the questionnaire is designed to be applied in both rural and urban areas, some of the questions are more relevant for one type of area than the other. Wherever possible, community informants should be asked screening questions that enable interviewers to skip sections on topics that are not relevant in the informants' communities. It is definitely preferable to devise a questionnaire that can be implemented in both types of areas with slight modifications than to have two different questionnaires; in practice, communities tend to be located somewhere on a continuum from "rural" to "urban" rather than in one of two clear and distinct categories. Using one of two possible questionnaires in a given community may result in the loss of important information. For example, in periurban locations, agriculture may be an important source of livelihood that should not be ignored in the community questionnaire, despite the fact that the location has been classified as urban.

Another reason why it is better to administer one questionnaire than two is that if two questionnaires are developed, it is easy to end up with information that is not comparable, which complicates the analysis of the resulting data set considerably. If certain sections of the questionnaire suggested here are relevant only for a certain type of community, it is better to

design and implement a set of screening questions than to rely on the crude urban-rural distinction to categorize communities.

Throughout the questionnaire, the word "community" is used in the same sense that it has been used throughout this chapter. Yet in each survey the term to be used may be different. In a survey that covers only rural areas, "village" may be the most appropriate term, and in socialist countries "commune" may be the best choice. In addition, the term "currency" should always be replaced by the name of the national currency, such as dollars, pesos, or rupees.

In defining response categories, the proposed questionnaire covers the most likely responses. However, in many countries survey designers should add new categories and delete some existing categories.

### **Annotations on the Questions in the Community Questionnaire**

This section explains parts of the draft questionnaire introduced in the previous section (and presented in Volume 3), detailing why particular choices were made and where variations may be appropriate.

#### ***Section 1: Cover Sheet and Respondent Characteristics***

The cover sheet collects basic information on the name and location of the community and on the identity of the interviewer, supervisor, and data entry operator. The section called "Respondent Characteristics" collects basic information about the people who participated in the interview. Instructions about how to select respondents, which will vary by country, can be added at the top as a "Note to the Interviewer." (Such instructions might include the minimum number of group members or the fact that a group should include one elected official and at least one teacher.) See also Chapter 4 on metadata.

#### ***Section 2: Physical and Demographic Characteristics of the Community***

It is useful to obtain information on basic physical and demographic characteristics of the community, such as land type (for example, coastal, inland, or mountainous), land use patterns, rainfall patterns (months of the rainy season), and geographical size—and on the number of households, individuals, and ethnic groups in the community. This information can be collected from community informants, but in the case of demo-

graphic characteristics, census data should also be obtained wherever possible. Because the primary sampling unit for most previous LSMS surveys has been a census enumeration area, census data usually exist, although they may be outdated, of questionable quality, or difficult to obtain. Questions may be added regarding kinship networks and the status of women. If these topics are of sufficient interest to the survey designers, an independent section should be created.

A specific point to bear in mind in this section is that, where possible, the survey team should obtain multiple Global Positioning System readouts on the latitude, longitude, and altitude of the community, so that average values can be used to minimize inaccuracy in these measures.

### *Section 3: Transportation*

The welfare of individuals is affected by their access to opportunities (for example, in labor markets, credit, education, and health and family planning services) and to information about these opportunities. Transportation networks increase access to these opportunities because they facilitate the transfer of goods, services, and information and enable individuals to respond to the availability of the goods, services, and information.

The transportation module in a community survey should include questions about the surface of the main roadway in the community and of the majority of community roads. Questions should also be included about how accessible the community is in different seasons and about whether a transportation service operates within the community and beyond the community. The questions should make it clear that a transportation service can be government-sponsored or privately operated and may or may not charge a fee. This module is also a good place to inquire about the price of fuel. Some items in this module may not be relevant in urban areas; it may therefore be appropriate to include an instruction to the interviewer that reads: "If the cluster is located in a busy urban area where public transportation is common and standard municipal services are available, skip to Question X." However, information on fuel prices should be collected in all communities.

It is also useful to inquire about the accessibility (within the community and through transportation services) of miscellaneous locations and services about which not much detail is needed, such as post offices,

telephone offices, and a daily market—as well as of administrative centers such as district and province capitals. The design of the transportation section of the community survey should take into account the types of services that are of special interest in the overall household survey. Certain services, such as schools, health and family planning facilities, and credit sources, are probably important enough to warrant their own modules. The questions about these services that appear in the transportation module should cover whether each service exists within the community, how far away the service is, and the travel cost and time required to reach the service by public transportation.

### *Employment in Industry (Section 4) and Agriculture (Section 5)*

A key measure of welfare is income (or consumption). Employment in the labor market is the primary source of income for most households. Consequently, in order to understand the determinants of economic welfare, it is necessary to know what labor market opportunities are available to the individuals living in a given community. These opportunities can be in both formal and informal sectors and can be in the agriculture sector, in small businesses, or in large industries. A basic employment module in the community questionnaire should yield information on the structure of the labor force in the community, including which sectors are present and which are of primary importance. This module should also inquire about wage rates for men, women, and children in agriculture, large industries, and small (cottage) industries. It should apply to both urban and rural communities, although not all questions need to be answered in all areas.

At minimum, the questions about agricultural employment should establish the important crops grown in the area and the wage rates for the different jobs associated with producing these crops. In many communities in the developing world, agriculture may be of such importance that this section should include detailed questions about how crops are produced, such as the number of harvests per year and the availability of key inputs like irrigation capacity or harvesting machinery. If land tenure is an important determinant of welfare, questions about this should also be included. Questions on agricultural extension activities can be included if they are relevant to the policy priorities of the overall survey. Questions about the importance

of animal husbandry in the community are also worthwhile.

The questions on employment in the industrial sector can follow a similar vein, beginning by inquiring whether there are large industries in the area. If so, determining the products and wages of the three nearest (or perhaps largest) factories can help characterize employment opportunities in this subsector.

Small (cottage) industries are another important source of employment in many developing countries. This subsector will be described by a small grid that determines whether such industries are an important source of employment in the community and, if so, what the industries produce and the wage rates the industries provide.

In some countries, public work schemes or unemployment insurance may be available to individuals who have had difficulty finding a job. If these programs provide an important social safety net, questions should be included about their availability in the community and what wages and benefits they offer.

Survey designers may wish to make the categories in Section 4, Questions 32–35 more detailed.

#### **Section 6: Credit**

The credit section in the draft module is fairly short and can be expanded depending on the focus of the household survey. If survey designers want a more complete set of credit institutions, a service availability roster can be used. (See Sections 9 and 10 on schools and health facilities.)

The initial questions focus on methods of saving and sources of credit. Additional questions ask for a list of the main places where community residents save or borrow money. For each of three places, questions are asked about the types of loans and methods of savings that are available, as well as the distance from the community center to the institution. The proposed module contains questions on interest rates. However, in places where interest rates vary depending on the characteristics of the borrower, these questions may be so difficult to answer that they should be deleted.

#### **Section 7: Physical Infrastructure**

Access to infrastructure is often used as a measure of welfare. It is clear that access to a safe water supply, for example, is a basic human need. The availability of infrastructure at the community level does not guarantee that the households in the sample survey use the

infrastructure, but it increases the likelihood that they do. Also, the availability of infrastructure at the community level may provide externalities to households (by contributing to a healthier environment) even if the household members do not use a particular service. The infrastructure module should focus on water sources, types of toilets, methods of sewage disposal, methods of solid waste disposal, and availability of electricity. For each of these categories of infrastructure, survey designers should consider including questions on:

- Major sources and methods.
- A source or method that accounts for most households in the community.
- Whether sources or methods vary seasonally and in a way deleterious to household welfare (for example, water shortages during the dry season).
- Whether services are publicly or privately provided or both.
- The extent to which services are disrupted.
- The year when a service became available.
- The current price for hooking up to or enrolling in a service.
- The proportion of households that use the “desirable” source or method (for example, piped water rather than pond water) or the source or method that has been a target of a government program (for example, the construction of sewage systems or the expansion of garbage collection services).

In asking questions about infrastructure, if the focus is on the local neighborhoods in which the survey households are located as opposed to on a larger community, it may not be necessary to ask questions about the proportion of households that use various methods because this information can be obtained by aggregating the household data (see earlier discussion).

#### **Section 8: History and Development**

One aspect of studying behavior involves studying how it changes in response to events that are beyond the control of individuals. These events can include natural disasters such as floods and droughts, changes in employment opportunities such as the opening of a new factory, and policy interventions such as the construction of public housing or a new clinic. The community questionnaire should include a module identifying events of this nature that have occurred during the previous 5–10 years. Data should be collected on the type of event, the year in which it occurred, and,

perhaps, an estimate of the proportion of the community affected.

Apart from describing the types of events that are of interest to survey designers, it is best to leave this section open-ended, allowing informants to decide what constitutes an important event. The reported events may overlap to some degree with other questions, but collecting small amounts of redundant information is preferable to omitting important changes. It is particularly important for this section that some of the informants in the group have lived in the community for some length of time.

### **Sections 9 and 10: Availability of Health Facilities and Schools**

Four types of services are crucial in most household surveys: health services, family planning services, schools, and credit sources. In some countries, health and family planning services are integrated so questions on the availability of both kinds of services can be combined. Three or four modules should be designed for the service availability section.

The first step is to determine, within the four broad categories, which types of facilities are relevant. For example, among educational institutions, the availability of universities and vocational academies is probably of less interest in most surveys than the availability of primary and secondary schools. Among health facilities, the basic government sources of primary care and family planning are likely to be particularly important in many countries, as are single-practice private providers. If the health of women and children is of interest to the survey designers, it will be important to include midwives in the list of service providers, while hospitals and traditional practitioners may be of less interest. The next step is to determine whether it is necessary to find out specific details about particular facilities in each category and, if so, how many facilities of each type are of interest.

A basic module on service availability might ask only about the number of facilities of different types in a specific geographical or administrative area. A more detailed module would obtain a list, by name and address, of the public and private services available to community members, as well as some information about the accessibility of each service provider on the list. Such a detailed list allows survey designers greater flexibility in determining the area that contains the services. (The same geographical or administrative

boundary need not apply to all areas.) It also has the advantage that it provides a list of facilities for which administrative data might be available, and, if a facility survey is done, it provides a list of facilities from which a sample can be drawn. Sections 9 and 10 reflect this detailed approach. If facility-specific identification codes are to be assigned (to facilitate linking household data to data in the community or facility survey), the service availability modules should include space for these identification codes.

When designing questions on service availability, it is important to keep in mind that the amount of detail that community informants can reliably provide about facilities is likely to be limited. The kind of information that they are likely to know includes whether the facility is within the administrative boundaries, the distance from the community center to the facility, and the time and expense required to travel to the facility. It may not be possible for community informants to summarize the full range of facilities available to everybody in the community. However, the information will become more complete as the size of the group of informants increases, because a wider range of experiences will be captured.

It may also be useful to ask about outreach programs that carry services into the community, such as immunization or family planning campaigns.

If a separate facility survey is conducted, the service availability modules can be abridged.

### **Direct Observation (Section 11)**

Interviewers necessarily spend several days in the community in which they are administering the household questionnaire. Survey designers can capitalize on this fact by including a section in which the interviewers record their observations of community features such as housing characteristics, environmental quality, security, and socioeconomic status of community members. (See Chapter 25 on qualitative data collection.) These observations by the interviewer will supplement the data on the demographic and physical characteristics of the community gathered from the community informants.

Survey designers may also be interested in gathering information—in this case metadata—from the survey team on its experiences in implementing and administering the survey. Useful questions might cover:

- The number of miles between each cluster and the previous cluster visited.

- The number of nights the team spent in a cluster.
- The address, price, and characteristics of the base camp.
- Whether electricity was available for entering data into computers.
- The names of community residents who were particularly helpful.

### Prices

Two price questionnaires are presented: one to be administered to community informants and one to be used at markets or with vendors. To minimize the

problem of missing values for prices, it is recommended that both questionnaires be used. The questionnaires are designed so that a preferred quantity and quality are specified for each item. The interviewers should first try to obtain a price estimate that corresponds to the specified quantity and quality. If the informant cannot provide that price, the interviewer should record a price that the informant can provide and the quantity and quality associated with that price. For more information see the subsection on measuring prices in the third section of this chapter.

### Appendix 13.1 Potential Items for the Price Questionnaire

Item	Examples of preferred units	Common units that are too vague	Examples of quality-related specifications
Alcohol	l liter	Bottle	Type, brand
Bananas	Kilogram, bunch of 6	Bunch	Variety
Beef	l kilogram		Which cut; boneless or with bones; cubes, strips, and other divisions; prepackaged or not
Beer	l liter	Bottle; can	Brand
Bread	l kilogram	l-loaf bag	Flour type, brand name
Cabbage	l kilogram	Head	
Cassava	l kilogram		Fresh roots, dried chips, flour variety
Chicken	l kilogram; if whole specify an approximate weight		Whole; plucked or unplucked; gutted, head and feet removed; parts (breasts, thighs, wings, necks, back, mixed); prepackaged or not
Cigarettes	Pack of 20; carton (10 packs of 20)	Pack	Brand; filtered or unfiltered
Cooking oil	l liter	Bottle	Source of oil (vegetable, animal, corn, safflower); brand
Corn	l kilogram		
Dried Beans	l kilogram	Bag	Variety (navy, pinto, black-eyed, and so on)
Duck			See chicken
Eggs	l egg; 6 eggs; 12 eggs; l kilogram		Chicken, duck, or other; refrigerated or not; how fresh (if this affects price)
Freshwater fish			See sea fish
Mangoes	l kilogram		Whole; variety
Milk (powdered)	l kilogram	Box	Brand
Milk (sweetened condensed)	Liter, milliliter	Can	
Millet	l kilogram		
Noodles		Package box	In bulk or brand name
Oranges	l kilogram	Bag	Variety
Papaya	l kilogram		Whole; variety
Peanuts		Bag; box	Shelled or unshelled; raw or roasted; salted or unsalted; with or without skins
Pork			See beef
Rice	l kilogram	Bag	Imported or exported; long-grained or short-grained; white, brown, or black; sticky rice hulled or unhulled; variety (local, modern, high-yielding)
Salt	500 grams	Box	Sea salt; iodized; level of refinement
Salted fish	l kilogram		
Sea fish	l kilogram; specify approximate weight if one fish is likely to weigh more than l kilogram		Whole; scaled or not; gutted or not; head removed or not; name of fish; fileted
Smoked fish	l kilogram		

**Appendix 13.1 Potential Items for the Price Questionnaire (continued)**

Item	Examples of preferred units	Common units that are too vague	Examples of quality-related specifications
Spinach		Bunch	Washed or not
Sugar	1 kilogram	Bag Box	Processed; level of refinement (description of granularity); type (white, brown, and so on)
Tobacco (loose)	500 grams	Bag	Brand
Tomatoes	1 tomato; 1 kilogram		Variety
Wheat flour	1 kilogram	Bag	Bulk or brand-name

Source: Author's summary.

**Nonfood Items**

- Aluminum saucepan
- Aspirin
- Bar soap
- Bicycle tire
- Bottled natural gas
- Cassette player
- Charcoal
- Coal dust
- Condoms
- Cotton cloth
- Firewood
- Fuel oil
- Iron rod
- Kerosene
- Laundry soap
- Light bulb
- Mosquito nets
- Oral contraceptives
- Oral rehydration solution
- Porcelain bowl
- Radio
- Rubber flip-flops
- Synthetic cloth
- Tin bowl
- Toothpaste

**Notes**

The author would like to thank Jere Behrman, Parfait Eloundou-Enyegue, Paul Glewwe, Margaret Grosh, Courtney Harold, John Strauss, Paramita Sudharto and Duncan Thomas for comments on earlier drafts of this chapter.

1. Sometimes, as was the case in the Ghana LSMS, households in a cluster span two communities. In this case interviewers should collect community data separately for each segment of the cluster that corresponds to a distinct community. Household questionnaires and community questionnaires must be labeled in the field so that they can be properly matched during data processing. Sometimes, as was the case in the Indonesian Family Life Survey,

two clusters are contained within the same community. In this case community data should be collected once and assigned to two distinct clusters, but facility data should be collected twice, depending on the protocol for selecting the facilities visited. Regardless of how neatly communities are defined, special cases will arise in the field. The best approach in such cases is to try to maintain consistency with the selected definition of a community.

2. This may not be true everywhere. In some countries, such as Ghana and Côte d'Ivoire, an entire administrative unit, such as a village, may serve as the cluster and thus have a great deal of significance.

3. The administrative structures of Vietnam and Indonesia are examples. In Vietnam the smallest unit is a village, and several villages comprise a commune, which is the lowest administrative level of government. In Indonesia households belong to "household groups," neighborhoods, villages, subdistricts, districts, and provinces, regardless of whether they are located in cities or in rural areas.

4. An analogy to the household questionnaire would be that the time unit to which expenditure questions refer varies by good according to assumptions about the frequency of purchase.

5. Four of the more recent LSMS surveys—Morocco, Nepal, Bolivia, and Ecuador—have included urban questionnaires. Initially, a similar strategy was pursued with the Service Availability modules in many of the Demographic and Health Surveys.

6. For example, administrators may believe that they will receive new equipment if they report that existing equipment malfunctions.

7. Matching two data sets that are based on the same geographic coding system is much easier than matching data sets based on different coding systems. If secondary data come from the same agency that is conducting the LSMS survey, it is more likely that the same coding system will be used than if two different agencies are involved.

8. Secondary data from government ministries often excludes private facilities, and community informants are often more knowledgeable about public facilities than about private facilities.

9. One can also obtain readings of latitude and longitude for the households, and thus obtain household-specific measures of distances to facilities. However, there are confidentiality issues with respect to releasing the coordinates of the households.

## References

- Beegle, K. 1995. *The Quality and Availability of Family Planning Services and Contraceptive Use in Tanzania*. Living Standards Measurement Study Working Paper 114. Washington, D.C.: World Bank.
- Entwistle, B., R. Rindfuss, S. Walsh, T. Evans, and S. Curran. 1997. "Geographic Information Systems, Spatial Network Analysis, and Contraceptive Choice." *Demography* 34 (2): 171–87.
- Frankenberg, E. 1995. "The Effects of Access to Health Care on Infant Mortality in Indonesia." *Health Transitions Review* 5 (2): 143–62.
- . 1998. "RAND's Experience with Collecting Facility Data." Paper presented at the Population Association of America meetings. Chicago, April.
- Feyisetan, B., and M. Ainsworth. 1994. *Contraceptive Use and the Quality, Price, and Availability of Family Planning in Nigeria*. Living Standards Measurement Study Working Paper 108. Washington, D.C.: World Bank.
- Frankenberg, E., and K. Beegle. 1998. "The Role of Government and Community in Women's Knowledge of Health Services." RAND, Santa Monica, Cal.
- Frankenberg, E., and L. Karoly. 1995. "Indonesia Family Life Survey 1993: Overview and Field Report." RAND. Publication Number: DRU-1195-NICHD/AID Santa Monica, Cal.
- Frankenberg, E., D. Thomas, and K. Beegle. 1999. "The Real Costs of Indonesia's Economic Crisis: Preliminary Findings from the Indonesia Family Life Surveys." Labor and Population Working Paper. DRU-2064-NIA/NICHD.
- Frankenberg, E., K. Beegle, B. Sikoki, and D. Thomas. 1998. "Health, Family Planning, and Well-being in Indonesia during an Economic Crisis." RAND Labor and Population Working Paper. Santa Monica, Cal.
- Gertler, P., and Molyneaux, J. 1994. "How Economic Development and Family Planning Programs Combined to Reduce Indonesian Fertility." *Demography* 31 (1): 33–64.
- Glewwe, P., and M. Grosh. 1998. "Data Watch: The World Bank's LSMS Household Surveys." *Journal of Economic Perspectives* 12 (1): 187–96.
- Glewwe, P., and H. Jacoby. 1994. "Student Achievement and Schooling Choice in Low-income Countries: Evidence from Ghana." *Journal of Human Resources* 29 (3): 843–64.
- . 1998. "School Enrollment and Completion." In David Dollar, Paul Glewwe, and Jennie Litvack, eds., *Household Welfare and VietNam's Transition to a Market Economy*. Washington, D.C.: World Bank.
- Glewwe, P., M. Grosh, H. Jacoby, and M. Lockheed. 1995. "An Eclectic Approach to Estimating the Determinants of Achievement in Jamaican Primary Education." *World Bank Economic Review* 9 (2): 231–58.
- Lavy, V. 1996. "School Supply Constraints and Children's Educational Outcomes in Rural Ghana." *Journal of Development Economics* 51 (2): 291–314.
- Lavy, V., J. Strauss, D. Thomas, and P. de Vreyer. 1996. "Quality of Health Care, Survival and Health Outcomes in Ghana." *Journal of Health Economics* 15 (3): 333–57.
- McCaffrey, Daniel. 1995. Conversation. October.
- Ravallion, M., and B. Bidani. 1992. "How Robust is a Poverty Profile?" World Bank, Washington, D.C.
- Rosenzweig, M., and K. Wolpin. 1986. "Evaluating the Effects of Optimally Distributed Public Programs: Child Health and Family Planning Interventions." *American Economic Review* 76 (3): 470–82.
- Rosenzweig, M., and K. Wolpin. 1988. "Migration Selectivity and the Effects of Public Programs." *Journal of Public Economics* 37: 265–89.
- Thomas, D., V. Lavy, and J. Strauss. 1996. "Public Policy and Anthropometric Outcomes in Côte d'Ivoire." *Journal of Public Economics* 61 (2): 155–92.
- van de Walle, D. 1995. *Infrastructure and Poverty in Vietnam*. Living Standards Measurement Study Working Paper 121. Washington, D.C.: World Bank.
- Vivjerberg, W. 1998. "Nonfarm Enterprises in Vietnam." In David Dollar, Paul Glewwe, and Jennie Litvack, eds., *Household Welfare and VietNam's Transition to a Market Economy*. Washington, D.C.: World Bank.





**The World Bank**