THE WORLD BANK ECONOMIC REVIEW

Volume 14 January 2000 Number 1

ISSN 0258-6770

Public Disclosure Authorized

THE WORLD BANK ECONOMIC REVIEW

The World Bank Economic Review is a professional journal for the dissemination of World Bank–sponsored research that informs policy analyses and choices. It is directed to an international readership among economists and social scientists in government, business, and international agencies, as well as in universities and development research institutions. The *Review* emphasizes policy relevance and operational aspects of economics, rather than primarily theoretical and methodological issues. It is intended for readers familiar with economic theory and analysis but not necessarily proficient in advanced mathematical or econometric techniques. Articles will illustrate how professional research can shed light on policy choices. Inconsistency with Bank policy will not be grounds for rejection of an article.

Articles will be drawn primarily from work conducted by World Bank staff and consultants. Before being accepted for publication by the Editorial Board, all articles are reviewed by two referees who are not members of the Bank's staff and one World Bank staff member; articles must also be recommended by at least one external member of the Editorial Board.

The *Review* may on occasion publish articles on specified topics by non-Bank contributors. Any reader interested in preparing such an article is invited to submit a proposal of not more than two pages in length to the Editor.

The views and interpretations expressed in this journal are those of the authors and do not necessarily represent the views and policies of the World Bank or of its 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 whatsoever for any consequences of their use. When maps are used, the boundaries, denominations, and other information 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.

Comments or brief notes responding to *Review* articles are welcome and will be considered for publication to the extent that space permits. Please direct all editorial correspondence to the Editor, *The World Bank Economic Review*, The World Bank, Washington, D.C. 20433, U.S.A.

The World Bank Economic Review is published three times a year (January, May, and September) by the World Bank. Single copies may be purchased at $12.95. Subscription rates are as follows:

	Individuals	*Institutions*
1-year subscription	US$30	US$50
2-year subscription	US$55	US$95
3-year subscription	US$70	US$130

Orders should be sent to: World Bank Publications, Box 7247–7956, Philadelphia, PA 19170–7956, U.S.A. Subscriptions are available without charge to readers with mailing addresses in developing countries and in socialist economies in transition. Written request is required every two years to renew such subscriptions.

Manufactured in the United States of America
ISBN 0-8213-4607-5; ISSN 0258-6770

This journal is indexed regularly in *Current Contents/Social & Behavioral Sciences, Index to International Statistics, Journal of Economic Literature, Public Affairs Information Service, and Social Sciences Citation Index*®. It is available in microform through University Microfilms, Inc., 300 North Zeeb Road, Ann Arbor, MI 48106, U.S.A.

THE WORLD BANK ECONOMIC REVIEW

Volume 14 January 2000 Number 1

A SYMPOSIUM ON TRADE AND TECHNOLOGY

A SYMPOSIUM ON GEOGRAPHICAL TARGETING

THE WORLD BANK ECONOMIC REVIEW, VOL. 14, NO. 1: 1-15

International Knowledge Flows and Economic Performance: A Review of the Evidence

Giorgio Barba Navaretti and David G. Tarr

An empirical analysis of the microeconomic links between trade and knowledge diffusion is useful for singling out some of the key predictions of the theory of endogenous growth in open economies. This literature postulates that total factor productivity is higher when trade gives countries access to a wider or more sophisticated range of technologies. The articles reviewed here find considerable evidence that imported technologies raise total factor productivity in importing countries, particularly developing countries and particularly when technologies are acquired by way of imports of intermediate goods. They also provide some support for the argument that exports and foreign direct investment are channels for learning. Although access to foreign technologies has a positive impact on developing countries' total factor productivity, overall these countries are shown to purchase older and simpler machines than industrial countries. Relative factor and machinery costs and skill and technology endowments affect the choice of imported technologies. However, government attempts to limit or guide the selection of technologies are likely to have a negative effect on growth because they discourage producers from purchasing the most appropriate and efficient machines. Rather, policies aimed at promoting technological development should strengthen the absorptive capacity of importing countries and address the complementarity between human and physical capital in a broader context.

In both developing and industrial countries there is an increasing institutional awareness of the importance of knowledge for business performance, economic growth, and development. The comparison of Ghana and the Republic of Korea is a tale of disparate growth that is frequently told. In 1960 both countries had the same income per capita; today Korea's is seven times higher. More staggering than the gap in performance is the inability of analysts to single out its causes.

Giorgio Barbara Navaretti is with the Università degli Studi di Ancona, Centro Studi Luca d'Agliano, Fondazione Eni Enrico Mattei, and David G. Tarr is with the Development Research Group at the World Bank. Their e-mail addresses are *barba@mailserver.unimi.it* and *dtarr@worldbank.org*. This symposium issue collects some of the papers presented at the conference on Trade and Technology Diffusion: The Evidence with Implications for Developing Countries, which was held in Milan on April 18–19, 1997, and was jointly organized by Fondazione Eni Enrico Mattei and the Development Research Group of the World Bank. The authors would like to thank participants at the conference for their comments. They especially thank James Tybout for his input in organizing the conference and for his extensive comments on this article.

The accumulation of physical and human capital barely justifies a threefold difference in income per capita. The remaining $4,000 gap is yet to be explained. Many analysts agree that knowledge could be the hidden factor of production that drove Korea's growth (World Bank 1999 and Rodriguez-Clare 1997).

The characteristics of knowledge as an economic good may explain why this initially neglected and now actively investigated factor of production remains relatively concealed. Knowledge is a public good with property rights that are rarely enforceable. It is seldom quantified or priced; it is sometimes codified, but more frequently tacit; and in any case it is difficult or impossible to observe. All measures of knowledge are indirect, either inputs to (years of schooling, manuals) or outputs of (human capital, patents, the unexplained residual in growth accounting) its accumulation.

Two elements combine to reveal why technology is likely to be important for economic growth: countries with higher per capita incomes also hold larger stocks of measured knowledge, and knowledge travels cheaply and fast. Although well established in theory and anecdotally, the link between growth and the accumulation of knowledge has been tested by relatively few rigorous empirical studies. The most quoted reference in this area is Coe and Helpman (1995), who focus on knowledge diffusion among Organisation for Economic Co-operation and Development (OECD) countries. Coe, Helpman, and Hoffmaister (1997) extend this work to developing economies.

Both studies use aggregate data to measure the impact of knowledge diffusion through trade flows. However, most of the effects of learning on productivity are observable primarily at the micro-level. For example, learning by doing and imported technologies are frequently sector-specific. Typically, technology is upgraded and production processes are enhanced at the firm level, guided by firms' characteristics. The organization and effectiveness of "knowledge departments" in corporations are also observable at the firm level. It is therefore important to combine the empirical analysis of generalized countrywide learning processes with a more focused understanding of how these processes take place at the micro-level.

This is precisely the objective of the articles collected in this symposium, which are devoted to the empirical analysis of the international diffusion of technological knowledge and economic performance at the micro-level. They analyze three international channels of learning: imports of machines and other inputs, along with the research and development (R&D) spillovers they generate; foreign direct investment (FDI) or other forms of collaboration between local and foreign firms; and learning by exporting.

If technology transfers are assumed to be exogenous, the scope of micro-level empirical work is simply to measure accurately the different components of this external learning process and to determine if this process does indeed work in the predicted direction. It is clear, however, that the international demand for and supply of technological inputs are not exogenous. Firms in developing countries will choose the type of technology imported so as to maximize their profits. Also,

firms in industrial economies choose the channel through which to transfer technologies based on profit-maximizing conditions. The form of the transfer (subsidiaries wholly owned by foreign investors, joint ventures, licensing, or direct sales of machinery) affects the dispersion of proprietary knowledge and thus monopoly rents. Working at the micro-level allows us to analyze the choice of the type of technology imported and the channel through which it is imported, along with its impact on economic performance.

I. Trade, Learning, and Growth: The Theoretical Background

The empirical studies collected here are rooted in theories developed by earlier models of endogenous growth in open economies. This literature recognizes two main mechanisms of knowledge accumulation. The first is that trade may change a country's pattern of specialization. Learning is faster if the country specializes in goods with higher learning potential (Krugman 1987; Lucas 1993; Stokey 1988, 1991; and Young 1991). The second mechanism is that trade in goods and factors of production may open new sources of technological inputs (Grossman and Helpman 1991 and Rivera-Batiz and Romer 1991). Broadly speaking, we can therefore consider two groups of models, those in which learning is essentially a domestic affair and those in which knowledge is imported from abroad.[1]

Many of the domestic learning models involve learning by doing. Young (1991), for example, develops a model in which managers and workers acquire experience through the manufacturing process that makes them more productive.[2] This well-documented phenomenon is typically summarized by a learning curve relating process-specific production costs to cumulative units produced.[3] The learning curve is downward-sloping, but it eventually flattens out as the potential for learning is exhausted. If each process is associated with a given product, the accumulation of knowledge is a nondecreasing, concave, and bounded function of the cumulative output of that product. Stokey (1991) shows in a second model that schooling can be substituted for learning by doing without significantly changing Young's (1991) conclusions.

Both Young (1991) and Stokey (1991) attribute productivity growth to learning processes that enable the production of increasingly sophisticated products and to the associated knowledge spillovers, which are limited to the country in which learning takes place. The more rapidly learning takes place—either through schooling

1. We do not provide a comprehensive survey but simply convey the main findings of this type of literature. For a comprehensive survey on trade and technology diffusion, see Grossman and Helpman (1995).

2. Earlier contributions involving growth and learning by doing include Lucas (1988), which elaborates on Krugman (1987), and Stokey (1988).

3. Many studies have found corroborating evidence that production experience lowers unit costs. Benkard (1997) makes an excellent contribution. He documents partial spillovers from experience in producing one generation of wide-body aircraft, making production in the next generation more efficient. Malerba (1992) provides a recent review of the literature and some evidence of his own.

or through learning by doing—the higher is the rate at which new high-end products are introduced, and the higher is the rate of productivity growth.

But if learning only takes place at home, why should we be concerned with trade? We should be concerned because trade affects the pattern of specialization in goods with different potential and indirectly influences the accumulation of knowledge. The domestic learning models imply that a likely, but not a necessary, consequence of trade liberalization in developing countries is the dampening of demand for high-end goods, thereby limiting the amount of learning.

Although the models of Young and Stokey are too restrictive to derive policy implications for developing countries, they are important contributions because they show clearly what happens when trade does not generate international knowledge flows. Indeed, not all learning is a by-product of learning by doing or of education. Knowledge diffuses across national boundaries in many ways. A country's knowledge may increase because its trading partners have accumulated knowledge, that is, there are international knowledge spillovers that are likely increased through trade.

Therefore, we now turn to models in which knowledge is not contained within national boundaries, and the channels through which it is transmitted are important. Trade (especially the import of a large variety of technological inputs), foreign investment, and travel on the part of executives and technicians are all channels through which knowledge can be transmitted internationally. Open economies benefit from an even wider range of learning channels.

One important model that allows for the international transmission of knowledge is that of Rivera-Batiz and Romer (1991). They assume an economy in which inputs are differentiated horizontally. Research creates new input varieties, and these new varieties raise productivity growth in the final-goods industry. When more input varieties are available, the final goods sector can choose inputs that more closely fit its precise requirements. Thus, for the same cost, it can purchase a more productive bundle of intermediate inputs. This is a dynamic extension of Ethier's (1982) application of the model of Dixit and Stiglitz (1977). Rivera-Batiz and Romer examine two channels for transferring technological knowledge: the transmission of ideas, which can be traded independently from goods—the knowledge-driven R&D model initially developed by Romer (1990)—and the trade of intermediate inputs that incorporate new ideas—the lab equipment model.

In both models international knowledge flows raise growth, as the stock of knowledge available to producers in any country increases. Note, though, that the two models imply very different mechanisms of technology diffusion. In the first case the stock of knowledge is common to all researchers in the world economy and is distributed at no cost by way of spillovers.[4] In the second case knowledge is immediately excludable in that it is only transmitted through the purchase of inputs, and there are otherwise no spillovers.

4. However, once a design is incorporated in an intermediate good, producers must pay for exclusive use of the design. This mechanism generates monopoly profits for the research sector and an incentive to innovate.

In the lab equipment model knowledge is diffused internationally only if there is trade in goods, since knowledge is embodied in goods. Trade improves productivity directly, and the steady-state growth rate increases. In the knowledge-driven model, if trade is permitted, countries have an incentive to specialize in completely non-overlapping innovations. Then knowledge diffuses through both trade in goods and the exchange of ideas. The welfare effect is larger since countries benefit from both the increase in productivity in manufacturing arising from additional foreign varieties and the increase in productivity in research arising from additional foreign ideas. Thus the form of knowledge diffusion is important.

Rivera-Batiz and Romer focus on symmetric countries. Grossman and Helpman (1991) extend this framework to the case of asymmetric countries and the case in which there is more than one final good.[5] In Grossman and Helpman's framework inputs are differentiated horizontally, and additional varieties developed by the research sector raise the productivity of the final-goods sector. The research technology is similar to that in Rivera-Batiz and Romer's knowledge-driven economy in that productivity depends on cumulative R&D stocks. An open country can use world R&D experience: if spillovers are global in scope, foreign R&D will have the same effect on productivity as domestic R&D. The same conclusions can be reached with a quality-ladder model, in which inputs are differentiated vertically.[6]

Grossman and Helpman make one prediction analogous to that of Young (1991).[7] Industrial economies with a relative abundance of human capital will undertake more research and grow faster than developing economies. But by engaging in international trade with industrial countries, developing countries can obtain a greater variety of intermediate inputs, and therefore grow faster, than they would otherwise. Although growth rates do not converge, openness does raise the growth rates of developing countries.

Thus, according to the endogenous growth literature, the impact of trade on the growth of developing countries depends crucially on the international scope of knowledge diffusion and on the mechanisms through which knowledge is transmitted. A way to test these predictions is to analyze how inflows of technologies affect productivity in the importing countries.

II. Imported Inputs, Learning, and Growth: The Empirical Evidence

Although earlier work estimated international knowledge spillovers, Coe and Helpman (1995) make the first and most widely quoted attempt to establish an empirical connection between international R&D spillovers and economic

5. Grossman and Helpman (1991) also include a model with no spillovers.

6. On quality ladders see also Aghion and Howitt (1992) and Segerstrom, Anant, and Dinopoulos (1990).

7. The model in Grossman and Helpman (1991) is essentially a Hecksher-Olin model of international trade combined with a model of endogenous growth through profit-seeking R&D. Young (1991), like Krugman (1987), uses a Ricardian model in which trade is driven by differences in technology rather than differences in comparative advantage. See chapter 11 in Aghion and Howitt (1998) for a survey of these models.

growth.[8] Coe and Helpman estimate how much of the variation in total factor productivity (TFP) for a sample of OECD countries is explained by the variation in domestic and foreign R&D capital stocks. Foreign R&D capital stocks are defined as the import share–weighted average of trade partners' domestic R&D capital stocks. The impact of foreign R&D in an importing country is expected to be higher the more the importing country buys from the foreign country undertaking the R&D.

Note the mechanisms of knowledge transmission implied by this specification. Knowledge is acquired through the purchase of intermediates, thus it does not travel if there is no trade in goods. Additionally, productivity in country i is expected to increase if the R&D capital stock in country j increases, even if j's share in i's total imports remains constant. One way to interpret this last assumption is that if the overall R&D stock in a given country grows, the goods imported from that country will be more R&D-intensive. Another interpretation is that, through trade, country i gains access to the R&D of country j, enabling country i to introduce new designs in production more easily. Coe and Helpman's (1995) empirical specification reflects Rivera-Batiz and Romer's lab equipment model in the first case and their knowledge-driven model in the second case.

Coe and Helpman find evidence that both foreign and domestic R&D improve TFP. Whereas in large countries the elasticity of TFP with respect to domestic R&D capital stocks is larger than that with respect to foreign R&D capital stocks, the opposite holds in small countries; that is, foreign R&D is more important for small countries. Coe, Helpman, and Hoffmaister (1997) confirm these results in their analysis of North-South R&D diffusion, based on a sample of 77 developing countries. They find that East Asian countries have benefited the most from foreign R&D. These results imply that access to foreign R&D is crucial for the typical developing country, and openness to and trade with industrial countries are fundamental to obtaining that access.

These two studies focus on aggregate TFP at the country level and assume that knowledge spillovers are channeled through total import flows. Yet much of the learning process is probably related to trade within industries and between firms.[9] Learning and technical progress mostly take place within specific activities.[10] This is certainly true for learning by doing, and it may also be true when new knowledge is embodied in specific machines or procedures.

Moreover, the potential for technical progress differs across industries. Consider two trade partners of country i—j and k—that have equal R&D capital stocks and equal weights in i's total imports, but differ in the basket of products that i imports from them. If the products imported from j are mostly consumer goods, whereas the products imported from k are mostly intermediate goods, we

8. For example, Jaffe, Henderson, and Trajtenberg (1993) examine geographic localization of knowledge spillovers by looking at patent citations. Eaton and Kortum (1996) analyze patterns of productivity and international patenting.

9. Hakura and Jaumotte (1999) rigorously support this hypothesis.

10. For a pioneering case study see Lall (1987).

would expect imports from *k* to have a greater impact on *i*'s TFP. Likewise, if most of the imports from *k* are goods with high learning potential relative to imports from *j*, *k*'s products will have a larger effect on *i*'s TFP.

The article by Wolfgang Keller in this symposium addresses this problem, examining the impact of international R&D spillovers on industry-specific TFP. His measure of foreign R&D for a given country is the trade-weighted sector-specific R&D capital stocks of its foreign partner countries. Keller computes trade weights by considering only imports of machines used in production in a given industry.[11] In doing so, he leans closer to the lab equipment model, because he focuses on imports that are more likely to embody new knowledge. His results are similar to those of Coe and Helpman (1995) in that domestic and foreign R&D stocks are estimated to have a significant and positive influence on TFP.

Keller also tests for the robustness of the trade weights used by carrying out an alternative battery of estimations using random weights for foreign R&D stocks. In line with Coe and Helpman's results he finds that domestic and foreign intermediates have a different impact and that this difference is statistically significant. Also, the impact of foreign intermediate imports is greater the smaller is the importing country. In contrast, he finds that the regression results are somewhat invariant to the import weights of trading partners' R&D stocks, especially when importing countries are large, when the trade shares of partners are evenly distributed, and when partners have similar R&D stocks. Consequently, trade is a more important conveyor of foreign technological knowledge to small countries than to large ones.

Since Keller studies only OECD countries, the question arises as to the implications of his results for developing countries. His finding that imported inputs are particularly effective for small OECD countries is important in this context. Developing countries typically have a weak domestic R&D sector and mainly acquire technologies internationally. For this reason it is likely that a change in import shares assigning a larger weight to developing economies, which have smaller R&D stocks, would clearly diminish their access to foreign technologies. In other words, the likely effect of a developing country's shift in acquiring machinery from Germany rather than from France is likely to be smaller than a shift in acquiring machinery from Ethiopia rather than from France.

Coe and Hoffmaister (1999) argue that Keller's random weights are in fact not random, but simple averages with a random error. They derive three alternative sets of random weights and with them restore fully their original results. However, this controversy concerns mostly technological flows between large countries, because Keller never dismisses the evidence that trade flows to small countries are important in transmitting foreign knowledge.[12] This is also consistent with the findings of Eaton and Kortum (1996). On the basis of patent data

11. Coe, Helpman, and Hoffmaister (1997) also use imports of intermediates, not total imports, to construct import shares.

12. See also the estimates for developing countries in Coe, Helpman, and Hoffmeister (1997).

they argue that 90 percent of growth in small OECD countries derives from foreign innovations.

In conclusion, the available empirical evidence supports the direct and indirect role of trade in diffusing knowledge and suggests that it is particularly important for developing countries to trade with technologically rich countries. This result is not surprising, given that rich countries have many more means of exchanging knowledge between themselves than developing countries. The articles reviewed so far also provide indirect evidence that channels other than foreign trade are at work, which may not be observed in aggregate estimations.

III. Learning at the Firm Level: Foreign Direct Investment and Exports

There is now fairly well documented evidence that multinational enterprises are a major channel for transferring technologies. According to UNCTAD (1997), more than 80 percent of royalty payments for international technology transfers were made from subsidiaries to their parent firm.[13] This is not surprising, given that multinationals are more often found in industries with high R&D expenditures relative to sales, a larger number of scientists and technicians, new and technically complex products, and high levels of product differentiation and advertising.[14] These stylized facts are well rooted in theory: intangible assets like knowledge are more mobile than tangible assets and are semi-public goods in the sense that knowledge used in one application does not reduce its availability for other applications either within or across firms.

The fact that multinationals transfer technologies does not necessarily imply that such transfers are beneficial for the local economy. It is necessary to determine whether the productivity of foreign-owned companies is higher than that of domestic companies and whether foreign-owned companies transfer knowledge to domestic companies. These spillovers can take place through many channels, including the movement of employees from foreign affiliates to domestic companies, backward and forward links between multinationals and domestic firms, and demonstration effects.

Early statistical analyses of spillovers at the industry level use a production function framework to estimate the impact on labor productivity of FDI (measured by the foreign share of each industry's employment or value added).[15] They all find significant evidence of positive spillovers.

More recent studies use panel data at the firm level for a few developing countries. Their conclusions are controversial. They report that firms with foreign ownership have higher TFP. In contrast, there is no evidence of positive short-run spillovers to domestic firms, and the concentration of foreign investment in particular sectors sometimes lowers the productivity of domestic firms in the same

13. This figure is reported in Saggi (1999), who extensively reviews this evidence. Blomström and Kokko (1998) offer another major survey of the role of multinationals in diffusing technology.

14. Markusen (1995) and Caves (1996) survey the theoretical and empirical literature on this matter.

15. Caves (1974), Globerman (1979), and Blomström and Persson (1983).

industry.[16] These results are partly explained by the fact that foreign investors acquire market shares at the expense of domestic producers, which face negative scale effects. Overall, the effects of FDI depend heavily on the absorptive capacity and the competitiveness of local firms. Spillovers will be larger if local firms are able to quickly adopt new imported technologies and to face the competition posed by more efficient foreign producers.[17]

The article by Simeon Djankov and Bernard Hoekman in this issue examines the impact of foreign investment on productivity for a sample of 513 Czech enterprises listed in the Prague stock exchange between 1992 and 1997. Their sample includes firms with foreign links (firms that had formed joint ventures or whose equity was at least partially foreign-owned) and firms without such formal links. In line with earlier contributions, Djankov and Hoekman examine the impact of formal foreign links on TFP and whether the presence of foreign investors generates positive indirect intraindustry spillovers on firms that have no foreign links.

As expected, the authors find evidence that firms with foreign links have higher TFP growth rates than firms without foreign links. Like some of the earlier studies using firm-level data, Djankov and Hoekman find that foreign investment lowers the productivity growth of domestic firms that have no formal foreign links. This result may reflect a competitive effect, as local firms without foreign links lose market share to those with foreign links. Firms without foreign links are still able to restructure and improve their productivity, but at a slower pace than their foreign-owned counterparts and at a slower pace than if there were no foreign investment in the industry.

Exporting is another channel through which firms based in open economies can acquire foreign knowledge. By learning about foreign markets, technologies, and products, they may specialize in products with high learning potential. Both case studies and empirical evidence support this view, showing that exporting firms are more efficient than nonexporting firms (Pack 1993). Yet this evidence says little about the causal relationship between firms' exporting status and their productivity. It may be that firms become more productive because they learn by exporting or it may be that more productive firms enter the export market in the first place. Clerides, Lach, and Tybout (1998) test the relevance of these two hypotheses for Colombia, Mexico, and Morocco, and Bernard and Jensen (1999) test them for the United States. Both studies find that self-selection is the dominant explanation and that there is little evidence of learning by exporting.

In this issue Bee Yan Aw, Sukkyun Chung, and Mark Roberts address this question for the Republic of Korea and Taiwan (China), using panels of exporting and nonexporting firms. They study the relationship between productivity and transitions into and out of the export market. If self-selection is the impor-

16. Harrison (1996); Aitken, Hanson, and Harrison (1997); and Haddad and Harrison (1993).

17. A key element missing from these studies, however, is an examination of the impact on downstream industries and overall economic welfare or growth. For example, if there is FDI in accounting services, although the domestic accounting sector may be adversely affected, productivity and welfare in the economy may rise. On this point see also Lall (1999).

tant explanation, then a firms' initial productivity should be higher when they enter the export market than firms that stay out. If learning by exporting is the relevant explanation, then producers who enter the market should experience higher productivity growth than firms that stay out.

For firms in Taiwan (China) Aw, Chung, and Roberts find evidence that both explanations are important. For Korea the results are much less satisfactory. There is weak evidence for the self-selection hypothesis and no evidence for the learning-by-exporting hypothesis. Aw, Chung, and Roberts claim that the weak results for Korea could be explained by the country's export policies, among other factors. Firms' decisions to enter the export market are linked more closely to their access to government promotion policies than to their ex ante productivity levels.

This discussion of Korea brings policy into the picture. It is clear that the link between international knowledge flows and productivity is affected by many factors, including policy. But to understand the role of policy we should briefly explore how the market for international knowledge flows works. Aw, Chung, and Roberts make endogenous the decision to export and, implicitly, the decision to learn. In the next section we explore the decisionmaking process characterizing other channels for acquiring knowledge.

IV. Markets for Imported Technologies and the Role of Policy

The empirical literature discussed above implicitly assumes that any country opening its borders to trade, even the most advanced country, will benefit from a wider variety of technologies and from technologies that are, at least in some fields, superior to those available in the domestic market. But the absorptive capacity of the liberalizing country is important in determining the input mix or the technologies imported.

We focus here on machines used in production, which represent the bulk of imported technologies. The choice of appropriate technology depends on the relative costs of the machines and on the costs of the factors of production needed to use them. Also, the availability of the required skills in a given firm or country is important in the choice of technology.[18] Skills, such as human capital or other technological capabilities, acquired through learning by doing or through formal training can be specific to a given technique. The greater is the complementarity between the required skills and the new technology, the more costly is the switch to an alternative technology.

The article by Giorgio Barba Navaretti, Isidro Soloaga, and Wendy Takacs in this symposium looks at the choice between new and used equipment when there are labor-saving technical progress and complementarity between technology and skills within the firm. Their analysis is based on a theoretical model of trade in used equipment among heterogeneous firms. If factor markets are imperfect, firms

18. See Benhabib and Rustichini (1991), Chari and Hopenhayn (1991), Jovanovic and MacDonald (1996), and Jovanovic and Nyarko (1995, 1998).

will face different factor prices even if they are based in the same country. The model allows firms to differ in their technical and managerial skills. Heterogeneity among firms located in different countries provides the underlying motive for trade in new and used machines and the basis on which the authors predict the share of used equipment imported.

They test these predictions by looking at U.S. exports of metalworking machine tools to 23 countries. Machines can be classified according to their vintage (new or used) and their technical complexity, which is measured in terms of the minimum skills necessary to use the machines. The empirical analysis shows that the share of used equipment is larger the lower is the per capita income of the importing country. Moreover, for low-income countries the share of used machines is larger the greater is the technical change embodied in new machines relative to old ones and the more complex are the skills required to use them efficiently. These results imply that technological factors and skill constraints are as important as relative factor prices in determining the choice of technology. They are thus consistent with results of other studies emphasizing that the ability of a given country to benefit from imported technologies depends on its absorptive capacity, that is, on its ability to master new and more complex technologies quickly and efficiently.[19]

These results also have important policy implications. First, policies designed to promote technological development should address the complementarities between different factors of production. For example, if there are explicit market failures in the education sector, fostering human capital should be a central policy goal. Second, policies that limit access to foreign capital—in particular, quantitative restrictions on imports of used machines—are ineffective and damaging. The aim of such policies is to foster technological upgrading. But, in fact, they deny firms access to more appropriate techniques and force them to buy machines that they may be unable to use efficiently.

This point is pursued further by David Guisselquist and Jean Marie Grether in the last article in the symposium. They analyze several case studies of technology transfer in agriculture, revealing that government attempts to select the right technologies to import in fact prevent local producers from gaining access to the best technologies. They identify two stylized patterns in the institutional arrangements governing the international flow of new agricultural inputs. Most economies, particularly industrial economies, rely on multiple channels in which farmers are exposed to new technologies because of the activities of different private and public actors. In contrast, many developing and transition economies centralize access to foreign technologies, making them subject to the approval of government committees based on lengthy performance tests.

Relying on two case studies in Bangladesh and in Turkey, Gisselquist and Grether show that deregulation leads to a significant increase in the transfer of technologies, a fall in the price of technologies, an increase in the use of new

19. See, for example, Nelson and Pack (1999) on the Asian Miracle.

technologies, and a dramatic increase in farmers' productivity. At least for seed technology, enough varieties were available worldwide to enable these developing countries to import and effectively use a subset of varieties to increase productivity.

V. Conclusions

The empirical analyses in this symposium allow us to assess some of the key predictions derived in the theoretical literature on endogenous growth. The articles reviewed here find considerable evidence that imported technologies raise TFP in the importing countries, particularly when technologies are acquired through imports of intermediate goods (Keller; Gisselquist and Grether). There is also some support for the view that exporting improves firm productivity (Aw, Chung, and Roberts). The role of FDI is more mixed. Although the productivity of the economy increases, foreign investment sometimes generates negative externalities on domestic producers in the same industry (Djankov and Hoekman).

Skill and knowledge endowments, along with relative costs, are shown to affect the choice of imported technologies. Developing countries are more likely to purchase older and simpler technologies the faster is the rate of technological progress embodied in new machines (Barba Navaretti, Soloaga, and Takacs).

These results bear important policy implications, especially for developing countries. The fact that imported technologies do not invariably carry a great potential for learning does not support government attempts to limit or guide the selection of technologies. Such policies hamper growth, because either they force producers to choose sophisticated technologies that they are unable to use or they prevent producers from buying the most appropriate and efficient technologies. The fundamental policy objective should be to allow as diverse a choice of technology inputs as possible, since diverse inputs are likely to increase productivity. Attempts to centralize decisionmaking and limit technology imports reduce productivity (Gisselquist and Grether).

Moreover, policies aimed at promoting technological development and at strengthening the absorptive capacity of importing countries should address the complementary relationship between human and physical capital. If the development of human capital is constrained by market failures in the education sector, technology policies should directly target such constraints. Indeed, all the theoretical literature and the evidence reviewed support the conclusion that inflows of technology are more beneficial the more quickly importers are able to master new and complex knowledge.[20]

The works discussed here can be extended in many directions. First, there is a clear need to better our understanding of learning processes at the micro-level.

20. Neary (1999) reviews research on R&D policy in developing countries and concludes that raising the general level of education is likely to have a more important impact than targeted policies like R&D subsidies.

We must define more accurate measures of technology flows and extend the collection of panel data at the firm level. Second, the literature on North-South technology flows generally assumes that the South purely imitates northern technologies. There is now growing evidence of dynamic R&D activity in many developing countries.[21] Further development of southern technologies could change the pattern of international knowledge flows. These developments should be studied more closely. Finally, innovations in information technology make it even more difficult to observe patterns of knowledge diffusion empirically. Thus empirical analysts attempting to identify international technology flows should be aware that the hidden factor of production they are searching for is becoming more and more concealed.

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth through Creative Destruction." *Econometrica* 60(2):323–51.

———. 1998. *Endogenous Growth Theory*. Cambridge, Mass.: MIT Press.

Aitken, Brian, Gordon Hanson, and Ann Harrison. 1997. "Spillovers, Foreign Investments, and Export Behavior." *Journal of International Economics* 43(1–2):103–32.

Barba Navaretti, Giorgio, and Carlo Carraro. 1999. "From Learning to Partnership: Multinational R&D Cooperation in Developing Countries." *Economics of Innovation and New Technologies* 8(1–2):137–73.

Benhabib, Jesse, and Aldo Rustichini. 1991. "Vintage Capital Investment and Growth." *Journal of Economic Theory* 55(December):323–39.

Benkard, Lanier. 1997. "Learning and Forgetting: The Dynamics of Aircraft Production." Yale University, New Haven, Conn. Processed.

Bernard, Andrew, and J. B. Jensen. 1999. "Exceptional Exporter Performance: Cause, Effect, or Both?" *Journal of International Economics* 47(1):1–25.

Blomström, Magnus, and Ari Kokko. 1998. "Foreign Investment as a Vehicle for International Technology Transfer." In Giorgio Barba Navaretti, Partha Dasgupta, Karl G. Mäler, and Domenico Siniscalco, eds., *Creation and Transfer of Knowledge: Institutions and Incentives*. Heidelberg and Berlin: Springer Verlag.

Blomström, Magnus, and Hakan Persson. 1983. "Foreign Investments and Spillover Efficiency in an Underdeveloped Economy: Evidence from the Mexican Manufacturing Industry." *World Development* 11(6):493–501.

Caves, Richard. 1974. "Multinational Firms, Competition, and Productivity in Host Country Markets." *Economica* 41(May):176–93.

———. 1996. *Multinational Enterprise and Economic Analysis*, 2d ed. London: Cambridge University Press.

Chari, V. V., and Hugo Hopenhayn. 1991. "Vintage Human Capital, Growth, and the Diffusion of New Technology." *Journal of Political Economy* 99(6):1142–65.

21. See Barba Navaretti and Carraro (1999).

Clerides, Sofronis, Saul Lach, and James Tybout. 1998. "Is 'Learning-by-Exporting' Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco." *Quarterly Journal of Economics* 114(3):452–75.

Coe, David, and Elhanan Helpman. 1995. "International R&D Spillovers." *European Economic Review* 39(5):859–87.

Coe, David, Elhanan Helpman, and Alexander Hoffmaister. 1997. "North-South R&D Spillovers." *Economic Journal* 107(January):134–49.

Coe, David, and Alexander Hoffmaister. 1999. "Are There International R&D Spillovers among Randomly Matched Trade Pairs?" IMF Working Paper 99/18. International Monetary Fund, Washington, D.C. Processed.

Dixit, Avinash, and Joseph E. Stiglitz. 1977. "Monopolistic Competition and Optimum Product Diversity." *American Economic Review* 67(3):297–308.

Eaton, Jonathan, and Samuel Kortum. 1996. "Trade in Ideas, Patenting, and Productivity in the OECD." *Journal of International Economics* 40(3–4):251–78.

Ethier, William. 1982. "National and International Returns to Scale in the Modern Theory of International Trade." *American Economic Review* 72(3):389–405.

Globerman, Steven. 1979. "Foreign Direct Investments and Spillover Efficiency Benefits in Canadian Manufacturing Industries." *Canadian Journal of Economics* 12(1):42–56.

Grossman, Gene, and Elhanan Helpman. 1991. *Innovation and Growth in the Global Economy.* Cambridge, Mass.: MIT Press.

———. 1995. "Technology and Trade." In Gene Grossman and Kenneth Rogoff, eds., *Handbook of International Economics.* Vol. 3. Amsterdam: Elsevier.

Haddad, Mona, and Ann Harrison. 1993. "Are There Positive Spillovers from Direct Foreign Investments? Evidence from Panel Data from Morocco." *Journal of Development Economics* 42(1):51–74.

Hakura, Dalia, and Florence Jaumotte. 1999. "The Role of Inter- and Intra-industry Trade in Technology Diffusion." IMF Working Paper 99/58. International Monetary Fund, Washington, D.C. Processed.

Harrison, Ann. 1996. "Determinants and Effects of Direct Foreign Investment in Côte d'Ivoire, Morocco, and Venezuela." In Mark Roberts and James Tybout, eds., *Industrial Evolution in Developing Countries.* New York: Oxford University Press.

Jaffe, Adam, Rebecca Henderson, and Manuel Trajtenberg. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics* 108(3):577–98.

Jovanovic, Boyan, and Glenn MacDonald. 1996. "Competitive Diffusion." *Journal of Political Economy* 102(1):24–52.

Jovanovic, Boyan, and Yaw Nyarko. 1995. "The Transfer of Human Capital." *Journal of Economic Dynamics and Control* 19(June):1033–64.

———. 1998. "How Does Productivity Depend on Past R&D?" In Giorgio Barba Navaretti, Partha Dasgupta, Karl G. Mäler, and Domenico Siniscalco, eds., *Creation and Transfer of Knowledge: Institutions and Incentives.* Heidelberg and Berlin: Springer Verlag.

Krugman, Paul. 1987. "The Narrow Moving Band, the Dutch Disease, and the Consequences of Mrs. Thatcher: Notes on Trade in the Presence of Scale Economies." *Journal of Development Economics* 27(1):41–55.

Lall, Sanjaya. 1987. *Learning to Industrialise: The Acquisition of Technological Capabilities in India.* London: Macmillan.

———. 1999. "Foreign Direct Investment, Technology, and Competitiveness: Trends and Challenges." Background paper prepared for United Nations Conference on Trade and Development, *World Investment Report 1999*. New York: United Nations.

Lucas, Robert. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22(1):3–42.

———. 1993. "Making a Miracle." *Econometrica* 61(2):251–72.

Malerba, Franco. 1992. "Learning by Firms and Incremental Technical Change." *Economic Journal* 102(413):845–59.

Markusen, James. 1995. "The Boundaries of Multinational Enterprises and the Theory of International Trade." *Journal of Economic Perspectives* 9(1):169–89.

Neary, Peter J. 1999. "R&D in Developing Countries: What Should Governments Do?" Department of Economics, University College, Dublin. Processed.

Nelson, Richard R., and Howard Pack. 1999. "The Asian Miracle and Modern Growth Theory." *Economic Journal* 109(45):416–36.

Pack, Howard. 1993. "Technology Gaps between Industrial and Developing Countries: Are There Dividends for Late-comers?" *Proceedings of the World Bank Annual Conference on Development Economics*. Washington, D.C.: World Bank.

Rivera-Batiz, Luis, and Paul M. Romer. 1991. "Economic Integration and Endogenous Growth." *Quarterly Journal of Economics* 106(2):531–55.

Rodriguez-Clare, Andres. 1997. "The Neoclassical Revival of Growth Economics: Has It Gone Too Far?" *Brookings 1997 Macroeconomics Annual*. Washington, D.C.: Brookings Institution.

Romer, Paul M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98(5):S71–102.

Saggi, Kamal. 1999. "Trade, Foreign Direct Investments, and International Technology Transfer." Development Research Group, World Bank, Washington, D.C. Processed.

Segerstrom, P. S., T. C. A. Anant, and Elias Dinopoulos. 1990. "A Schumpeterian Model of Product Life Cycle." *American Economic Review* 80(5):1077–91.

Stokey, Nancy. 1988. "Learning-by-Doing and the Introduction of New Goods." *Journal of Political Economy* 96(4):701–17.

———. 1991. "Human Capital, Product Quality, and Growth." *Quarterly Journal of Economics* 106(2):587–616.

UNCTAD (United Nations Conference on Trade and Development). 1997. *World Investment Report: Transnational Corporations, Market Structure, and Competition Policy*. New York: United Nations.

World Bank. 1999. *World Development Report 1999: Knowledge for Development*. New York: Oxford University Press.

Young, Alwyn. 1991. "Learning-by-Doing and the Dynamic Effects of International Trade." *Quarterly Journal of Economics* 106(2):369–405.

THE WORLD BANK ECONOMIC REVIEW, VOL. 14, NO. 1: 17-47

Do Trade Patterns and Technology Flows Affect Productivity Growth?

Wolfgang Keller

This article presents a model suggesting that the pattern of a country's intermediate goods imports affects its level of productivity because a country that imports such goods primarily from technological leaders receives more technology than a country that imports primarily from follower countries. The importance of trade patterns in determining technology flows is quantified using industry-level data for machinery goods imports and productivity from eight member countries of the Organisation for Economic Co-operation and Development between 1970 and 1991. Three conclusions emerge from this work. First, the eight countries studied appear to benefit more from domestic research and development (R&D) than from R&D of the average foreign country. Second, conditional on technology diffusion from domestic R&D, a country's import composition matters only if it is strongly biased toward or away from technological leaders. Third, differences in technology inflows related to the pattern of imports explain about 20 percent of the total variation in productivity growth. The implications of these findings for developing countries are discussed.

There is wide agreement among economists today that differences in physical and human capital accumulation alone do not explain the large variation in economic growth across countries. The important complementary role of technological diffusion in raising rates of economic growth has long been recognized, but little is known about the specific policies that promote such diffusion, particularly at the international level.

A widely held view is that international trade leads to faster technological diffusion and higher rates of productivity growth (Helpman 1997). Whereas these changes are important for all countries, they have dramatic implications for developing countries as they seek to catch up with the technological leaders in the Organisation for Economic Co-operation and Development (OECD). International

Wolfgang Keller is with the Department of Economics at the University of Texas, Austin and the National Bureau for Economic Research. His e-mail address is *keller@eco.utexas.edu*. An early version of this article was presented at the conference Trade and Technology Diffusion: The Evidence with Implications for Developing Countries, held in Milan, April 1997, jointly sponsored by the Fondazione Eni Enrico Mattei and the International Trade Division of the World Bank. The author thanks participants of the conference, as well as the late Zvi Griliches, Elhanan Helpman, Juergen Jerger, Sam Kortum, Dan Trefler, Jim Tybout, three referees, and the editor, who provided helpful criticism on an earlier draft. He also benefited from seminar presentations at Berlin (Humboldt), Bonn, Freiburg, Rochester, Rutgers, Tilburg, and Toronto.

agencies such as the World Bank routinely recommend policies that foster international trade, in part because it is presumed to further international technology diffusion (World Bank 1991, 1998). To date, however, there is little sound evidence to support this view.[1]

The recent development of theories of endogenous technological change, in particular the work of Romer (1990) and Aghion and Howitt (1992), has stimulated new analyses of the relationship among trade, growth, and technological change in open economies (Grossman and Helpman 1991; Rivera-Batiz and Romer 1991). In that work the authors embed the recent theories in general equilibrium models to analyze how trade in both intermediate and final goods affects long-run growth. Technology is diffused in this framework by being embodied in intermediate inputs: if research and development (R&D) expenditures create new intermediate goods that are different (the horizontally differentiated inputs model) or better (the quality ladder model) from those already existing and if these goods are also exported to other economies, then the importing country's productivity will increase through the R&D efforts of its trade partner.

The framework suggested by these models is well suited to studying empirically how trade patterns determine technology flows that trigger productivity growth and what impact importing a new (or better) type of intermediate product might have. First, the possibility of employing a larger range of intermediate inputs in output production allows for a productivity-enhancing increase in the degree of specialization in the production of intermediate inputs. To the extent that the importing country succeeds in not paying in full for this increase in variety, it reaps an external benefit, or, spillover effect. Second, the import of specialized inputs might facilitate learning about the product, spurring imitation or innovation of a competing product.

In this article I use data on the G-7 countries (Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States) and Sweden to evaluate these mechanisms. The traded goods studied are machinery inputs for manufacturing industries, which are usually differentiated and imperfect substitutes. Machinery inputs are also often highly specialized, implying that the elasticity of substitution between inputs produced for two different industries is negligible.

In this setting I examine whether productivity growth in an importing country is increased by the R&D investments of its trade partners. It is clear that the pattern of trade in intermediate inputs is central to this technology diffusion hypothesis. Both the increasing variety and the reverse-engineering effects are tied to arm's length market transactions of goods.[2]

The first hypothesis concerns the composition of imports by a partner country. On average, countries that import largely from high-knowledge countries should, everything else equal, import more and better-differentiated input varie-

1. See the review of the literature below, as well as Aghion and Howitt (1998) for a broader discussion.

2. Of course, other mechanisms of international technology diffusion, such as foreign direct investment (FDI), do not depend on imports. On FDI, see Lichtenberg and van Pottelsberghe de la Potterie (1998a) and Wang and Xu (1996).

ties than countries importing largely from low-knowledge countries. As a result, productivity should be higher in such countries than in countries that import from low-knowledge countries.

The second hypothesis posits that for a given composition of imports, this effect is likely to be stronger the higher is a country's overall import share. There are many reasons why a higher import share (or more openness) might lead to higher growth rates, however, some of which have little to do with international technology diffusion (see, for example, Edwards 1993). Because the claim that import composition matters is shared by fewer other models than the claim that a high overall level of imports is beneficial, I focus on import composition in order to address the question of whether trade is an important determinant of international technology diffusion. For the same reason examining the effect of import composition is a more powerful test of the recent trade and growth models than is studying the overall effect of imports.[3]

A number of recent articles have attempted to assess the importance of imports in transmitting foreign technology to domestic industries and spurring total factor productivity (TFP) growth (Coe and Helpman 1995; Coe, Helpman, and Hoffmaister 1997; Evenson 1995; Keller 1997, 1998a, 1998b; Lichtenberg and van Pottelsberghe de la Potterie 1998b).[4] Coe and Helpman (1995), who initiated this research, find a significant positive correlation between TFP levels and a weighted sum of partner-country R&D stocks, in which bilateral import shares serve as weights.[5]

The interpretation of this finding is not clear, however. Using the same data, Keller (1998a) finds that the role played by the composition of imports is limited: alternatively weighted R&D stocks—in which import shares are created randomly—also lead to a positive correlation between foreign R&D and the importing country's R&D. Moreover, the average correlation is often larger than when foreign R&D is weighted using observed import shares.[6]

While making the point that Coe and Helpman's (1995) correlations do not depend on the observed patterns of imports between countries, Keller's (1998a) results do not imply that R&D spillovers are unrelated to international trade, for several reasons. First, Coe and Helpman (1995) use aggregate import data to compute the trade-share weights for a given country. Overall import relations between countries, however, are likely to be a poor measure of relations in inter-

3. For further discussion of the implication of such model identification problems, see Evenett and Keller (1998).

4. Park (1995), Bernstein and Mohnen (1998), and Branstetter (1996) also estimate international R&D spillovers but do not include an explicit argument with respect to international trade. See also Eaton and Kortum (forthcoming).

5. Coe and Helpman (1995), as well as other authors, consider not only the effect of import composition but also the effect of technology inflows resulting from the level of imports, with the import composition given. Both might contribute to international technology diffusion. In this article, however, I focus primarily on the effect of import composition for model identification reasons.

6. Keller (1998a) and Coe, Helpman, and Hoffmaister (1997) also report regressions of TFP on unweighted R&D stocks, something that Coe and Helpman (1995) consider as well. The relationship of these regressions to the analysis in this article is discussed in section IV.

mediate goods trade. Second, most R&D is conducted in only a relatively small set of manufacturing industries (OECD 1991). Thus inferences that might be drawn from analysis of data at a more disaggregated level may not be possible based on country-level analysis. Third, common trends and shocks that affect R&D and productivity simultaneously might lead to spurious regression results that cloud any real international technology diffusion related to trade patterns. Finally, even if trade patterns are not the only determinants of international technology diffusion, it is necessary to quantify their contribution in order to assess the relative importance of such patterns.

In this article I analyze R&D, imports, and productivity at the two- and three-digit industry level. One is much more likely to observe trade flows embodying new technology at this level of aggregation than at the country level. I present estimation results for both TFP levels and TFP growth rates and address some of the open questions concerning common trends and simultaneity. (In appendix A I also report different sets of auxiliary regressions that analyze the robustness of the findings.) I also extend the Monte Carlo analysis of Keller (1998a), showing how such experiments are related to estimating an overall spillover effect from foreign R&D. This analysis allows me to determine whether some international R&D spillovers are related to trade patterns.

I. Theoretical Framework

The following gives a brief background of the recent theory that guides the empirical analysis performed here.[7] Long-run growth is endogenously determined by R&D investments, and technology is transmitted through trade in intermediate inputs. Assume that a country's output is produced according to

$$z = Al^{\alpha}d^{1-\alpha},\ 0 < \alpha < 1 \tag{1}$$

where A is a constant, l are labor services, and d is a composite input consisting of horizontally differentiated goods x of variety s:

$$d = \left(\int_0^{n^e} x(s)^{1-\alpha} ds\right)^{\frac{1}{1-\alpha}}. \tag{2}$$

The variable n^e denotes the range of intermediate inputs *employed* in a country; it can differ from n, the range of intermediate inputs *produced* in a country. The variable n increases when entrepreneurs devote resources to R&D (χ). If R&D capital does not depreciate, the range of intermediate inputs at time T will equal

$$n(T) = \int_{-\infty}^{T} \chi(t)dt \tag{3}$$

7. See Aghion and Howitt (1992), Grossman and Helpman (1991), Rivera-Batiz and Romer (1991), and Romer (1990). The books by Barro and Sala-i-Martin (1995) and Aghion and Howitt (1998) offer broader perspectives on the topic.

that is, the cumulative resources devoted to R&D up to time T. I define $n(T) \equiv S(T)$.

The goods $x(s)$ are best thought of as differentiated capital goods; they are produced with forgone consumption, or capital, denoted by k. Under certain conditions one can express the total intermediate input usage d in terms of capital k. Using equation 1, we can write the reduced-form expression for output as

$$z = A(n^e)^{\alpha} l^{\alpha} k^{1-\alpha}. \tag{4}$$

If F is TFP, defined as $F \equiv z / l^{\alpha} k^{1-\alpha}$, using equation 4:

$$\log F = \log A + \alpha \log n^e. \tag{5}$$

Equation 5 shows that productivity is positively related to the amount of variety of the employed product.

Many countries, $v = (i,h,...,V)$, will import foreign intermediates rather than use only domestic varieties, implying that each country employs a larger range of intermediate goods than it produces itself. In this sense the possibility of trade allows each country a greater degree of specialization in the production of intermediate goods than would be possible without trade. Specialization raises productivity, because the constant elasticity of substitution specification in equation 2 implies that for a given quantity of primary resources, output is increasing in the range of differentiated inputs (Ethier 1982). International trade leads to increases in productivity because only one country has to invent a new product variety (by spending the fixed R&D cost, χ), whereas *all* countries can potentially employ the new product by importing it.

When there are many countries and industries, denoted $j = (1,...,J)$, the composite input d of country i's industry j, d_{ij}, is given by

$$d_{ij} = \left(\int_0^{n^i_{ij}} x^i_{ij}(s)^{1-\alpha} ds + \gamma^i_{hj} \int_0^{n^i_{hj}} x^i_{hj}(s')^{1-\alpha} ds' + \ldots \right)^{\frac{1}{1-\alpha}}. \tag{6}$$

Here, $x^i_{ij}(s)$ denotes the quantity of an intermediate good of variety s used in sector j. The country that produces the intermediate good is given by the subscript; the superscript denotes the country that employs the intermediate good. Similarly, n^i_{ij} gives the range of domestically produced intermediate goods employed in country i's production of good j, and n^i_{hj} is the range of goods that country i imports from country h. The variable γ^i_{hj} determines the degree of substitutability between intermediates produced in country h and domestic intermediates produced in country i's industry j. If substitutability is perfect, then the γ's equal one.[8]

8. To capture the often highly specialized nature of machinery inputs for particular industries, I assume that only inputs of type j are productive in any country's sector j. Keller (1998b) also examines interindustry technology flows.

For simplicity, most theoretical work has concentrated on one-industry, two-country models (Rivera-Batiz and Romer 1991; Keller 1996).[9] As a result, empirical work using a multi-country, multi-industry setting has usually not estimated structural equations of such models of trade, technology diffusion, and growth. Instead, to go from the structural relationship of productivity and R&D in the one-sector closed-economy model (equation 5) to the multi-country, multi-industry context, researchers have related productivity to both domestic and foreign R&D, $\omega \neq v$:

$$F_{vj} = \Psi(n^v_{vj}, n^v_{wj}) = \Phi(S_{vj}, S_{wj}, \ldots), \forall v, j$$

where $\Psi(.)$ and $\Phi(.)$ are unknown functions.

However, the model of trade in intermediate inputs predicts that productivity in country v is related to R&D in country $w \neq v$ only to the extent that country v employs imported intermediates from country w. Productivity in country v should depend on country v's bilateral import share from w, denoted by m^v_w (the import composition effect), as well as country v's overall import share, denoted by m_v. At the industry level this means that

$$F_{vj} = \Phi(S_{vj}, S_{wj}, m_{vj}, m^v_{wj}), \forall v, j. \tag{7}$$

One can think of the import shares in equation 7 as indicating the probability of receiving a new type of foreign intermediate. This is certainly the correct interpretation in the extreme case in which $m^v_w = 0$. In all other cases, however, there is not necessarily a link between the level of imports and the number of types of newly introduced intermediate goods in the local economy.[10] Grossman and Helpman (1991) suggest several reasons why it is likely that the number of new varieties adopted from a partner country is positively related to the import volume from that country. This assumption guides the empirical specifications in section III.

II. Data

I examine data for eight OECD countries—Canada, France, Germany, Italy, Japan, Sweden, the United Kingdom, and the United States—in six sectors for the years 1970–91. (See appendixes B and C for data sources and the construction of the variables.) These sections include International Standard Industrial Classification (ISIC) 31 (food, beverages, and tobacco); ISIC 32 (textiles, apparel, and

9. Exceptions include Grossman and Helpman (1990), Feenstra (1996), and Aghion and Howitt (1998).

10. Especially if one also considers indirect effects, such as the possibility that importing leads to local learning through reverse-engineering and the subsequent invention of new inputs, it becomes clear that the volume of imports is an imperfect measure of the increase in varieties available domestically. An interesting alternative, albeit one with problems of its own, has been considered by Klenow and Rodriguez-Clare (1996), who postulate that the number of different varieties of intermediate goods is related to the number of different trade partners a country has.

leather); ISIC 341 (paper and paper products); ISIC 342 (printing); ISIC 36 and 37 (mineral products and basic metal industries); and ISIC 381 (metal products). All sectors belong to ISIC class 3 (manufacturing). The reliability and comparability of the measurement of inputs and outputs is high in these sectors relative to nonmanufacturing sectors.

The data on imports of machinery come from the OECD *Trade by Commodities* statistics (OECD 1980). I have tried to identify machinery imports that have a high probability of being used exclusively in one of the six manufacturing industries. These commodity classes are (revision 2) Standard International Trade Classification (SITC) 727 (food-processing machines and parts); SITC 724 (textile and leather machinery and parts); SITC 725 (paper and pulp mill machinery, machinery for manufacturing of paper); SITC 726 (printing and bookbinding machinery and parts); SITC 736 and 737 (machine tools for working metals and metal-working machinery and parts); and (revision 1) SITC 7184 and 7185 (mining, metal-crushing, and glass-working machinery). The bilateral trade relations for these SITC classes are given in full in tables A-1 to A-6 in appendix A.

OECD (1991) data on R&D expenditures by sector are used to capture the ranges of intermediate inputs, n. These data cover all intramural business enterprise expenditures on R&D. Because none of these industries has a ratio of R&D expenditures to value added of more than 0.5 percent, it is reasonable to assume that insofar as their productivity benefits from R&D at all, it will be largely from R&D performed outside the industry. Because there are no internationally comparable data on R&D in the machinery industry products used in specific industries, I assume that R&D expenditures toward sector j's machinery inputs are equal to a certain constant share of the R&D performed in the country's nonelectrical machinery sector (ISIC 382), where all specialized new machinery inputs are likely to be invented.[11] R&D stocks are derived from the R&D expenditure series using the perpetual inventory method. Descriptive statistics on the cumulative R&D stocks are given in table A-7.

I construct the TFP index using the *Structural Analysis Industrial* (STAN) database of the OECD (1994). The share parameter α is, by profit maximization of the producers, equal to the ratio of total labor to production costs. As emphasized by Hall (1990), using cost-based rather than revenue-based factor shares ensures robustness of the TFP index in the presence of imperfect competition, as in the model sketched above. Building on the integrated capital taxation model (see Jorgenson 1993 for an overview), I construct cost-based labor shares. The variables l, the number of workers employed, and y, gross production, come from the STAN database. The growth of the TFP index, F, is the difference between output growth and input growth weighted by factor-cost shares, with the level of the F's normalized to 100 in 1970 for each of the 8×6 time series. Summary statistics for the TFP data are shown in table A-8.

11. This constant share is industry employment divided by total manufacturing employment over the years 1979–81. The employment data are from OECD (1994).

III. Estimation Results

Below I present and discuss TFP-level estimation results. Then, I report and discuss estimation results for TFP growth rate regressions.

TFP-Level Specification

Consider the following specification:

$$\log F_{vjt} = \mu_j d_j + \delta_v d_v + \sum_{c \in G7S} \beta_c (m_{cj}^v \log S_{cjt}) + \varepsilon_{vjt}, \forall v, j, t. \tag{8}$$

The subscript t indexes a period, c indexes any of the eight countries in our sample (denoted as $G7S$ in equation 8), d_j represents industry fixed effects, and d_v represents country fixed effects. In this specification the TFP level in any industry is related to R&D in the same industry in all eight countries. The domestic weight, m_{vj}^v, $\forall v, j$, is set to one, and the weights of the partner countries are given by the bilateral import shares, $\sum_{w \neq v} m_{wj}^v = 1, \forall v, j$. The eight country elasticities, β_c, are constrained to be the same across importing countries.[12]

In equation 8 the import shares pick up differences in import composition across countries, which, according to the theory, affect the degree to which the importing country benefits from foreign technology. The specification also implies that two countries with the same import composition but different overall import shares benefit to the same degree from foreign R&D—an unlikely outcome if a higher overall import share increases a country's chance of benefiting from foreign technology. Following Coe and Helpman (1995), I model the contribution of a country's openness to imports for a given import composition by including the overall import share, m_{vj}, in the specification:[13]

$$\log F_{vjt} = \mu_j d_j + \delta_v d_v + \sum_{c \in G7S} \beta_c (m_{vj} m_{cj}^v \log S_{cjt}) + \varepsilon_{vjt}. \tag{9}$$

I refer to a specification that does not include the overall import share, as in equation 8, as NIS; a specification that includes the overall import share is referred to as IS.[14]

12. This approach differs from that of Coe and Helpman (1995), who estimate one parameter for the effect of foreign R&D. In addition, here, the bilateral import shares enter linearly, not logarithmically.

13. For the own-R&D effect, m_{vj} is chosen such that $m_{vj} m_{vj}^v \log S_{cjt} = \log S_{cjt}, \forall v, j, t,$ that is, m_{vj} then equals one.

14. Lichtenberg and van Pottelsberghe de la Potterie (1998b) criticize Coe and Helpman's (1995) specification using the overall import share because it implies an indexation bias. Their criticism does not apply here, because I have not indexed the R&D stocks. Lichtenberg and van Pottelsberghe de la Potterie (1998b) also point out that Coe and Helpman's weighting scheme suffers from a strong aggregation bias (country mergers or break-ups would strongly affect the estimated spillovers). I have not investigated this question in the present context, but Wang and Xu (1996) compare the weighting schemes proposed by Coe and Helpman (1995) and Lichtenberg and van Pottelsberghe de la Potterie (1998b), and propose a third scheme themselves.

Both specifications, equations 8 and 9, might be subject to simultaneity and omitted variable biases. These two problems would imply that the ordinary least squares (OLS) estimates are inconsistent, because the regressors and the error term are correlated. Since both productivity and R&D trend upward over time, part of the estimated correlation between the variables in equations 8 and 9 could be due to a common trend. In addition, the error could contain price or demand shocks that affect productivity and R&D jointly.[15]

To some extent, these problems can be reduced by imposing reasonable a priori restrictions on the dependent variable, TFP.[16] The importance of spurious correlation due to common trends can be assessed by comparing level regressions with results from growth regressions (presented in the next section), in which time-differencing eliminates common trends.[17]

Even with a growth specification, however, the possibility remains that exogenous shocks in the error term are correlated with changes in R&D activity. The solution to this problem is typically instrumental variable (IV) estimation. However, a standard choice for instrumenting quantity series—namely, real factor prices—is not available in the R&D context. Moreover, it is hard to obtain data on other variables that would serve as good instruments for cumulative R&D stocks for all countries, industries, and years in the sample.

If no good instruments are available, consistent parameters can still be estimated in the panel context by including a full set of fixed effects, provided the error can be decomposed into a permanent component that affects the regressor and a temporary component that does not. In this way, the part of the error that is correlated with the regressor will then be subsumed into the estimated fixed effect. Griliches and Hausman (1986) show that including a large number of fixed effects exacerbates errors-in-variables problems, however, which are also likely to be present in this context. The productivity-level specifications reported above together with the growth specifications reported below represent a compromise among these considerations. Additional auxiliary regressions are discussed in appendix E.

Results for the specifications given in equations 8 and 9 show that R&D stocks in all countries have a significant and positive influence on the TFP level of the receiving country (table 1). The magnitude of these effects, however, varies substantially. In the second specification, for example, results range from a low of 1.9 percent for Germany to a high of 27.6 percent for Sweden. The specifications account for one-third to one-half of the variation of the TFP indexes across countries, with the NIS specification yielding a higher R^2.

Do the results really say anything about the international diffusion of technology? To what extent, for example, do these results depend on correlating TFP

15. This section draws on Griliches (1979, 1995).

16. See appendix C for details.

17. For comparison purposes, in appendix E, I discuss TFP-level regressions that include time fixed effects.

Table 1. *Total Factor Productivity–Level Specifications*

Country	*NIS*[a]	*IS*[b]
Canada	0.101	0.201
	(0.027)	(0.043)
France	0.209	0.236
	(0.019)	(0.024)
Germany	0.071	0.019
	(0.009)	(0.009)
Italy	0.066	0.083
	(0.014)	(0.015)
Japan	0.068	0.127
	(0.014)	(0.020)
Sweden	0.206	0.276
	(0.022)	(0.025)
United Kingdom	0.188	0.150
	(0.022)	(0.027)
United States	0.111	0.080
	(0.007)	(0.011)
R^2	0.472	0.357

Note: All parameter values are statistically significant at the 5 percent level. Asymptotic standard errors are shown in parentheses. n = 1,056.

a. Specification does not include overall import share (see equation 8).

b. Specification does include overall import share (see equation 9).

with R&D, as opposed to physical capital? Perhaps technology is embodied in the physical capital stocks of the countries, and correlating TFP with foreign capital stocks would produce results similar to those found by correlating productivity with foreign R&D. To examine this question, I construct these physical capital stocks, denoted by K_{vjt}, and use them instead of the R&D stocks S_{vjt} in the specifications given by equations 8 and 9. The K_{vjt} variables are based on the estimated capital stocks in the nonelectrical machinery industries of the eight countries (ISIC 382); their construction is analogous to the R&D stocks S_{vjt}. In the NIS specification (equation 8) substituting K_{vjt} for S_{vjt} results in a drop in the R^2 from 0.472 to 0.169. In the IS specification (equation 9) the R^2 also falls substantially, from 0.357 to 0.179. Thus variation in R&D levels accounts for much more of the variation in TFP levels, suggesting that cumulative R&D captures the economies' stocks of technology better than physical capital does.

The result that high stocks of weighted foreign R&D are associated with high domestic productivity is interesting, but as such it does not say much about the importance of the fact that the weighting variables are the observed bilateral import shares. If these shares are interpreted as the probability that the importing country receives new intermediate inputs from a partner country, a natural question to ask is how the estimated parameters would look if we had employed a different set of probability weights, corresponding to different import patterns.

To examine this issue, I conduct Monte Carlo experiments. I investigate two different questions.[18] First, conditional on the effect of domestic R&D on productivity, is there evidence indicating that the composition of intermediate imports matters for productivity growth across sectors? Second, is there support for the hypothesis that foreign and domestic R&D have different effects on productivity?

DOES PRODUCTIVITY PERFORMANCE REFLECT THE COMPOSITION OF INTERMEDIATE IMPORTS? In the following experiments I randomly switch the bilateral import shares of given importing countries. Let b denote a specific Monte Carlo replication, $b = 1,\ldots,B$. The experiments are constrained such that only the composition of international demand is randomized. That is, the results are conditional on the domestic R&D effect: $\theta^{v}_{vj}(b) = 1$, $\forall v, j, b$. For all $w \neq v$ this means that

$$\theta^{v}_{wj}(b) = m^{v}_{qj} \text{ with } \Pr = \frac{1}{7},\ q \in G7S \setminus v, \forall v, w, j. \tag{10}$$

The $\theta^{v}_{wj}(b)$ are constructed such that $\sum_w \theta^{v}_{wj}(b) = 1$, that is, any observed import share is assigned only once.[19] The two specifications are

$$\log F_{vjt} = \mu_j d_j + \delta_v d_v + \sum_{c \in G7S} \beta_c [\theta^{v}_{cj}(b) \log S_{cjt}] + \varepsilon_{vjt}, \forall v, j, t, b \tag{11}$$

and

$$\log F_{vjt} = \mu_j d_j + \delta_v d_v + \sum_{c \in G7S} \beta_c [m_{vj} \theta^{v}_{cj}(b) \log S_{cjt}] + \varepsilon_{vjt}, \forall v, j, t, b. \tag{12}$$

In 75 percent of the cases the average estimates of β_c are significantly different from zero and positive (table 2). In addition, these coefficients are sometimes smaller and sometimes larger than those obtained by employing observed import shares. Thus no clear pattern can be detected. Moreover, the regressions that

18. Another approach to gauge the importance of the m^{v}_{cj} might be to simply drop them from equations 8 and 9. In equation 8 only two of eight β_c parameters are then estimated to be significantly different from zero (that of Japan, at –2.02, and that of the United States, at 1.37). If the bilateral import shares in equation 9 are dropped, only the β_c parameters for Germany, Italy, and Sweden are significantly positive; the parameter for France is significantly negative and those of the other countries are not significantly different from zero. Clearly, according to this test, the bilateral import shares matter. What is not clear so far, however, is whether the import composition or only the fact that the m^{v}_{cj} are not equal to one also matters.

19. For a given industry and importing country I draw seven numbers from a uniform distribution with support [0,1]. These are matched with the seven observed import shares to form a 7×2 matrix. This matrix is then sorted in ascending order on the random number column. In this way the probability that any import share $\sigma^{v}_{wj}(b)$ is equal to the value m^{v}_{wj}, for all w, is equal to 1/7. A new sequence of trade relations (the seven numbers from the uniform distribution with support [0,1]) is drawn for every importing country and every industry, making a total of $8 \times 6 = 48$ independent sequences.

Table 2. *Total Factor Productivity–Level Regressions*

	NIS			IS		
Country	*Observed shares*[a]	*Import shares switched*[b]	*Shares switched*[c]	*Observed shares*[d]	*Import shares switched*[e]	*Shares switched*[f]
Canada	0.101**	0.159	0.191	0.201**	0.104	0.026
	(0.027)	(0.081)	(0.097)	(0.043)	(0.085)	(0.253)
France	0.209**	0.161**	0.132	0.236**	0.180**	0.028
	(0.019)	(0.063)	(0.068)	(0.024)	(0.081)	(0.156)
Germany	0.071**	0.118**	0.115**	0.019**	0.128**	0.107
	(0.009)	(0.042)	(0.052)	(0.009)	(0.049)	(0.132)
Italy	0.066**	0.087**	0.134	0.083**	0.083**	0.243
	(0.014)	(0.028)	(0.080)	(0.015)	(0.028)	(0.308)
Japan	0.068**	0.103**	0.123**	0.127**	0.097**	0.034
	(0.014)	(0.043)	(0.053)	(0.020)	(0.046)	(0.136)
Sweden	0.206**	0.172**	0.147**	0.276**	0.253**	0.200
	(0.022)	(0.053)	(0.072)	(0.025)	(0.042)	(0.244)
United Kingdom	0.188**	0.134**	0.134	0.150**	0.165	0.028
	(0.022)	(0.064)	(0.067)	(0.027)	(0.086)	(0.129)
United States	0.111**	0.082**	0.108**	0.080**	0.081	0.035
	(0.007)	(0.039)	(0.043)	(0.011)	(0.044)	(0.092)
R^2	0.472	0.490	0.522	0.357	0.379	0.260

**Significant at the 5 percent level

Note: Standard errors are shown in parentheses. n = 1,056. The NIS specification does not include overall import share; the IS specification does include overall import share.

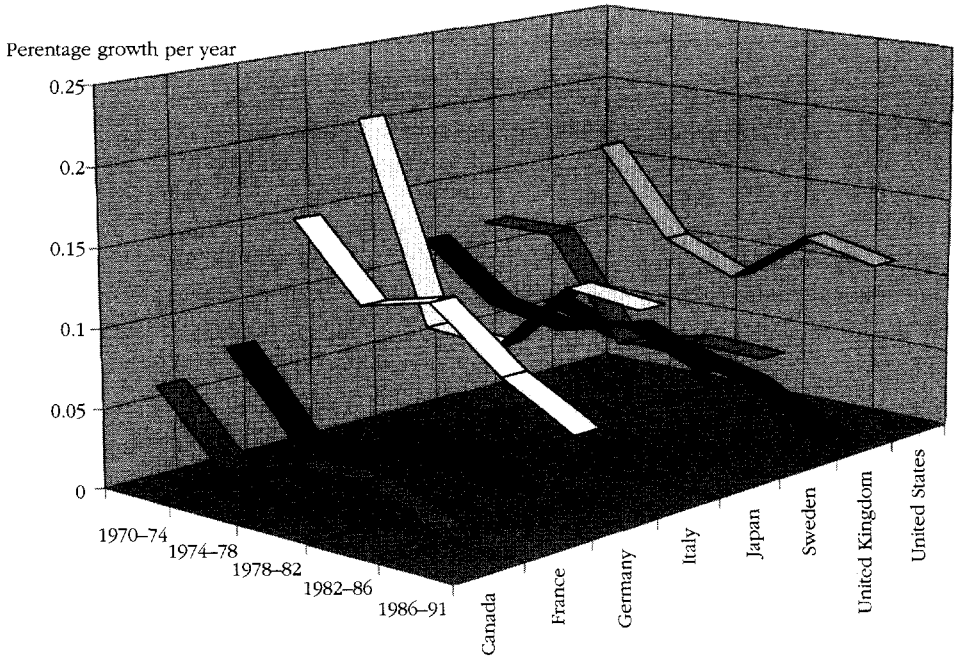
a. See equation 8.
b. See equation 11.
c. See equation 14.
d. See equation 9.
e. See equation 12.
f. See equation 15.

employ randomly switched import shares account for a slightly higher share of the variation in productivity than the observed import share regressions.

The fact that it is not necessary to impose the observed import shares to estimate significant international R&D spillovers parallels the finding of Keller (1998a) that one cannot test the hypothesis of links between R&D, the trade pattern, and TFP simply by examining whether the parameter estimates are positive or the R^2 is high. Obviously, the regression results are to some degree invariant to the weights with which the R&D stocks are multiplied. This would be trivial if the R&D stocks of different countries were equal in size and moved together over time. However, as shown in table A-7, there are considerable differences in the cumulative R&D stocks of different countries. In addition, the R&D stocks of different countries do not grow at the same rates, nor do they rise and fall simultaneously (figure 1).[20] This reasoning, at least in its extreme form, thus cannot explain the finding

20. The average annual growth rates of the R&D stock estimates range from 3.64 percent for Canada to 11.88 percent for Italy. The standard deviations of these growth rates for different four-year subperiods across countries range from a low of 2.87 percent (1978–82) to a high of 5.15 percent (1970–74).

Figure 1. *Rate of Growth of Machinery R&D Across Countries, 1970–91*



that the parameter estimates are to some extent invariant to switching the import shares.

Another interpretation of these results is that although import composition matters, conditional on the effect of domestic R&D, its impact is limited. It is important to realize that the effect of a country's import composition on its productivity is identified primarily from particularly low and particularly high import shares. Clearly, if all countries imported from their partners to the same extent, exchanging bilateral import shares would have no effect on the regression results.

This notion will be made more precise below. At this point, it is worth noting that the parameters for R&D from Canada and the United States are estimated very imprecisely and are often not different from zero at standard significance levels (table 2). This result is consistent with the idea that the trade-related effect of R&D is identified primarily from countries with extreme trade patterns, such as that exhibited by the United States and Canada, which have very large import shares with each other (tables A-1 to A-6). This trade pattern differs substantially from a symmetric trade pattern, in which countries import equal shares from all partners. Technology flows from these countries do not significantly affect productivity once the import shares are randomized.

FOREIGN AND DOMESTIC INTERMEDIATE INPUTS: DOES IT MATTER HOW MUCH AND FROM WHERE? In these experiments I switch the observed bilateral shares randomly,

including the weight on domestic R&D ($m^v_{vj} = 1, \forall v,j$). Any bilateral import share in replication b, $\sigma^v_{cj}(b)$ is thus equal to

$$\sigma^v_{cj}(b) = \begin{cases} m^v_{vj} \text{ with } \Pr = \frac{1}{8} \\ \quad\vdots \\ m^v_{Vj} \text{ with } \Pr = \frac{1}{8} \end{cases}, \forall v, c, j. \tag{13}$$

Because $m^v_{vj} = 1$ and $\sum_{w \neq v} m^v_{wj} = 1$, $\forall v,j$, it holds that $\sum_c \sigma^v_{cj}(b) = 2$, $\forall v,j$. Hence the experiment reveals whether, conditional on the value for $m^v_{vj} = 1$, $\forall v,j$ chosen ex ante, it is important to distinguish between embodied technology in intermediate inputs from domestic producers and from foreign producers. The equations are

$$\log F_{vjt} = \mu d_j + \delta d_v + \sum_{c \in G7S} \beta_c [\sigma^v_{cj}(b) \log S_{cjt}] + \varepsilon_{vjt}, \forall b, v, j, t \tag{14}$$

for the NIS specification and

$$\log F_{vjt} = \mu d_j + \delta d_v + \sum_{c \in G7S} \beta_c [m_{vj} \sigma^v_{cj}(b) \log S_{cjt}] + \varepsilon_{vjt}, \forall b, v, j, t \tag{15}$$

for the IS specification.

In 75 percent of the cases the Monte Carlo experiments result in coefficient estimates that are statistically indistinguishable from zero at the 5 percent level (table 2). For the model given by equation 14, half of the coefficient estimates are not significant; none of the β_c estimates for the model given by equation 15 is significant. The average R^2 in column three (0.522) is larger than that for the corresponding observed data regression. This finding is surprising, but could be spurious. Overall, the result that parameter estimates tend not to be significantly different from zero in the Monte Carlo experiment implies that if use of intermediate inputs, produced abroad or domestically, is determined randomly, the statistically significant relationship between R&D and productivity disappears.

To summarize, significant and quantitatively important productivity effects from R&D are found if the domestic source of technology diffusion is distinguished from foreign sources, while no robust relationship between R&D and productivity is found if domestic and foreign sources are treated symmetrically. It thus follows that the source of technology diffusion (domestic or foreign) matters. Moreover, because the domestic R&D weight, m^v_{vj}, is set to equal one, $\forall v,j$, whereas only the sum of foreign R&D weights equals one ($\sum_{w \neq v} m^v_{wj} = 1$, $\forall v,j$), the comparison of the observed share results and the randomized share results indicates that domestic R&D has a stronger impact on productivity than R&D from the average foreign country. This suggests that international technology diffusion might be nationally, or, more generally, geographically localized for these countries.

In appendix E I discuss some auxiliary regressions that include more fixed effects and a time trend in the basic specifications given by equations 8 and 9. Overall, the results in the appendix suggest that the findings above are robust.

Estimation of TFP Growth

The TFP growth specifications corresponding to equations 8 and 9 are

$$\frac{\Delta F_{vjt}}{F_{vjt}} = \alpha_0 + \sum_{c \in G7S} \beta_c \left(m_{cj}^{v} \frac{\Delta S_{cjt}}{S_{cjt}} \right) + \varepsilon_{vjt}, \forall v, j, t \tag{16}$$

where $\Delta x/x$ denotes the average annual growth rate of any variable x, and $m_{vj}^{v} = 1, \forall v, j$. The specification that includes the overall import share is given by

$$\frac{\Delta F_{vjt}}{F_{vjt}} = \alpha_0 + \sum_{c \in G7S} \beta_c \left(m_{vj} m_{cj}^{v} \frac{\Delta S_{cjt}}{S_{cjt}} \right) + \varepsilon_{vjt}, \forall v, j, t \tag{17}$$

where the value of the import share from v, m_{vj}, is set equal to one, $\forall v, j$. Dividing the period of observation into five subperiods of about four years each yields 240 observations (table 3).[21]

All slope coefficients are estimated to be positive. Moreover, in the IS model they are also significantly different from zero at the 5 percent level. The IS appears to be the preferred specification in this class of models. This result is consistent with the arguments given above as well as with the findings of Coe and Helpman (1995), even though the R^2 here is lower in the IS than in the NIS specification.

Only the results of the Monte Carlo experiments that were conditional on the effect of domestic R&D—those in which only the import shares were switched—are presented (table 3). The specifications are

$$\frac{\Delta F_{vjt}}{F_{vjt}} = \alpha_0 + \sum_{c \in G7S} \beta_c \left(\theta_{cj}^{v}(b) \frac{\Delta S_{cjt}}{S_{cjt}} \right) + \varepsilon_{vjt}, \forall b, v, j, t \tag{18}$$

and

$$\frac{\Delta F_{vjt}}{F_{vjt}} = \alpha_0 + \sum_{c \in G7S} \beta_c \left(m_{vj} \theta_{cj}^{v}(b) \frac{\Delta S_{cjt}}{S_{cjt}} \right) + \varepsilon_{vjt}, \forall b, v, j, t. \tag{19}$$

For each of these specifications I conduct 1,000 experiments. All Monte Carlo–based coefficients are estimated to be significantly greater than zero, confirming

21. The fact that the specifications given by equations 16 and 17 do not include industry or country fixed effects means that these industries are assumed to share a common growth rate. In reality, this assumption might be violated, and I have run some auxiliary regressions that include industry and country fixed effects into these growth regressions. In most cases these fixed effects are not estimated to be different from zero at standard levels of significance, so I do not include them here. See appendix E for more details.

Table 3. *Total Factor Productivity Growth Specifications*

	NIS		*IS*	
Country	*Observed shares*[a]	*Import shares switched*[b]	*Observed shares*[c]	*Import shares switched*[d]
Canada	0.351*	0.383**	0.415**	0.427**
	(0.178)	(0.122)	(0.158)	(0.022)
France	0.437**	0.431**	0.503**	0.512**
	(0.139)	(0.078)	(0.141)	(0.018)
Germany	0.198**	0.210**	0.235**	0.252**
	(0.067)	(0.027)	(0.060)	(0.009)
Italy	0.093*	0.126**	0.151**	0.157**
	(0.054)	(0.030)	(0.053)	(0.007)
Japan	0.068	0.077**	0.166**	0.169**
	(0.076)	(0.037)	(0.080)	(0.010)
Sweden	0.153**	0.155**	0.172**	0.189**
	(0.072)	(0.037)	(0.070)	(0.008)
United Kingdom	0.380**	0.358**	0.493**	0.508**
	(0.153)	(0.077)	(0.158)	(0.018)
United States	0.137**	0.108**	0.173**	0.173**
	(0.062)	(0.024)	(0.061)	(0.009)
R^2	0.127	0.134	0.105	0.109

*Significant at the 10 percent level

**Significant at the 5 percent level

Note: Standard errors are shown in parentheses. $n = 240$. The NIS specification does not include overall import share; the IS specification does include overall import share.

a. See equation 16.

b. See equation 18.

c. See equation 17.

d. See equation 19.

the earlier results from productivity-level regressions. Moreover, the mean estimates from the Monte Carlo experiments are very similar to the coefficients in the corresponding observed trade-share regression. For instance, a 95 percent confidence interval for the coefficient of Canada in the IS specification is equal to $0.427 \pm (2 \times 0.022)$. That this interval also includes the estimate for the import-weighted R&D effect from Canada when employing observed data (0.415), implies that the Canadian trade-related R&D effect is not statistically different from a randomized Canadian R&D effect, as captured by the average Monte Carlo estimate. The only parameters β_c that are significantly different in the randomized-share results compared with the observed-share results are for Sweden in the IS specification and Japan in the NIS specification.

IV. Separating Trade-Related R&D Spillovers from Average R&D Spillovers

In this section I show how switching the import shares is related to an average spillover effect from foreign R&D. I also examine whether the pattern of international trade influences international technology diffusion.

Monte Carlo Experiments and Average Foreign R&D Spillovers

Consider the average of a particular off-diagonal R&D weight across the B simulations, $\sigma_w^v(\bar{b}) = 1/B \sum_b \sigma_w^v(b)$, $\forall w \neq v$. Because the exchanging of the m_w^v is independent and identically distributed (i.i.d.) as $B \to \infty$, this average will be the same for all, $\sigma_w^v(\bar{b}) = \sigma(\bar{b}), \forall v,w$. With seven trade partners for any importing country, given that $7 \times \sigma(\bar{b}) = 1$, $\sigma(\bar{b}) = 1/7$.[22] Hence for any partner country's R&D variable across all B replications,

$$(20) \qquad \frac{1}{B}\sum_b \left(\sigma_w^v(b)\frac{\Delta S_{wj}}{S_{wj}}\right) = \frac{\Delta S_{wj}}{S_{wj}}\frac{\sum_b \sigma_w^v(b)}{B} = \sigma(\bar{b})\frac{\Delta S_{wj}}{S_{wj}}, \forall w \neq v.$$

Therefore, across all B replications the average regressors are the average annual growth rates, $\Delta S_{wj}/S_{wj}$, $w \neq v$, multiplied by $\sigma(\bar{b}) = 1/7$ for all partner countries and by $\Delta S_{vj}/S_{vj}$, the own-country R&D variable. Note, however, that the coefficients reported from the Monte Carlo experiments are averages across the OLS estimates from 1,000 replications, not OLS estimates from employing the average regressors. Nevertheless, as I show in appendix D, the two will be very similar under certain circumstances, both because the regression equation is linear and because the trade weights enter the specification linearly. The Monte Carlo–based estimates can then be viewed as estimating average R&D spillover effects. In table 4 I present the following average R&D spillover regression[23]

$$(21) \qquad \frac{\Delta F_{vjt}}{F_{vjt}} = \alpha_0 + \sum_{c \in G7S} \beta_c \left[m_{vj}\left(\sigma(\bar{b})\frac{\Delta S_{cjt}}{S_{cjt}}\right)\right] + \varepsilon_{vjt}.$$

Comparing this regression with the Monte Carlo–based results from table 3, it is clear that the Monte Carlo averages indeed estimate the average R&D spillover effect. The maximum relative difference between the estimated parameters is 2 percent (18.5 percent compared with 18.9 percent in the case of Sweden).[24]

Estimating the Contribution of Trade Patterns in Accounting for Productivity Growth across Countries

The previous section suggests a direct way of assessing whether there is a marginal international R&D spillover that is related to international trade patterns. Consider the following regression:

$$(22) \quad \frac{\Delta F_{vjt}}{F_{vjt}} = \alpha_0 + \sum_{c \in G7S} \beta_c^I \left[m_{vj}\sigma(\bar{b})\frac{\Delta S_{cjt}}{S_{cjt}}\right] + \sum_{c \in G7S} \beta_c^{II} \left[m_{vj}\left(m_{cj}^v - \sigma(\bar{b})\right)\frac{\Delta S_{cjt}}{S_{cjt}}\right] + \varepsilon_{vjt}.$$

22. This argument applies to any way of creating random import shares as long as it is i.i.d. and imposes $\sum_w \sigma_{wj}^v(b) = 1, \forall b,v$. It encompasses the procedure in Keller (1998a) and also randomizations that create arbitrary random shares, as opposed to randomization through exchanging the observed import shares, as has been done here.

23. Even if m_{vj} does not enter the following equation, the regression is feasible because the weight for domestic R&D is set equal to one. Otherwise, the average spillover regression (without m_{vj}) would not be feasible, because the regressor $\sigma(\bar{b})\Delta S_w/S_w$ would not vary by importing country. See the discussion in Lichtenberg (1993) on estimating the impact of a general R&D spillover effect that is the same for all countries.

24. The estimated standard deviations in these two regressions are not comparable.

Table 4. *Total Factory Productivity Growth Estimations*

Country	*Average R&D spillover*[a]	*Import shares switched*[b]
Canada	0.426	0.427
	(0.156)	(0.022)
France	0.513	0.512
	(0.139)	(0.018)
Germany	0.252	0.252
	(0.062)	(0.009)
Italy	0.156	0.157
	(0.052)	(0.007)
Japan	0.167	0.169
	(0.080)	(0.010)
Sweden	0.185	0.189
	(0.069)	(0.008)
United Kingdom	0.508	0.508
	(0.157)	(0.018)
United States	0.173	0.173
	(0.061)	(0.009)
R^2	0.109	0.109

Note: All values significant at the 5 percent level. Standard errors are shown in parentheses. $n = 240$.
a. See equation 21.
b. See equation 19.

The eight regressors with parameters β^{I} measure the average R&D spillover effect, and the eight β^{II} coefficients estimate the marginal trade pattern–related effect, if any. If there is no separate effect of international R&D that works through the pattern of international trade, then the coefficients β^{II} will be equal to zero, and equation 22 will explain as much of the variation in productivity growth rates as the average R&D spillover specification (equation 21) (table 5). The specification allowing for an additional R&D spillover effect related to trade patterns explains more of the variation in TFP growth rates than the specification that captures only the average R&D spillover effect (the adjusted R^2 values are 9.6 percent compared with 7.8 percent). The marginal effect of the bilateral trade pattern thus explains about 20 percent of the productivity growth effect from international R&D spillovers.[25]

The β^{II} point estimates in table 5 can be interpreted to indicate that industries that purchase a large share of their imports from a particular country (that is, more than 1/7 of total imports) experience on average lower rates of productivity growth than those that import the same share from all trading partners. The

25. Table 5 shows the adjusted R^2 as the number of regressors in the two specifications differs. I have considered analogous regressions for the NIS growth specification, as well as for the TFP-level NIS and IS regressions to check the robustness of this finding. I estimate that bilateral trade patterns accounted for 7.8 percent of the total international R&D spillover effects in the TFP-level NIS specification and 26.5 percent of the total in the IS specification. In these cases the restricted regression setting the β^{II} coefficients to zero is rejected at all standard levels of significance. In the NIS growth specification no significant marginal trade-related R&D spillover effect is estimated. Hence, while not perfectly robust, the pattern of bilateral trade is estimated to contribute significantly to understanding the total productivity effect from foreign R&D, accounting for about 20 percent of the effect in the preferred specification.

Table 5. *Total Factory Productivity Growth Estimations*

	Average spillover[a]	*Average and trade spillover*[b]	
Country	β_v	β^I	β^{II}
Canada	0.426**	0.389*	−13.61**
	(0.156)	(0.231)	(4.30)
France	0.513**	0.398**	4.11
	(0.139)	(0.181)	(3.39)
Germany	0.252**	0.126	−1.24
	(0.062)	(0.083)	(0.82)
Italy	0.156**	0.102*	−2.10
	(0.052)	(0.061)	(1.43)
Japan	0.167**	0.129	1.44
	(0.080)	(0.086)	(2.28)
Sweden	0.185**	0.165*	−1.22
	(0.069)	(0.090)	(2.19)
United Kingdom	0.508**	0.310	−5.12
	(0.157)	(0.189)	(6.19)
United States	0.173**	0.157**	−0.39
	(0.061)	(0.067)	(0.84)
Adjusted R^2	7.8	9.6	

* Significant at the 10 percent level.
** Significant at the 5 percent level.
Note: Standard errors are shown in parentheses. $n = 240$.

effect is estimated to be positive for France and Japan and negative for all other countries. It is significantly different from zero at the 5 percent level only for Canada, however.

V. Conclusions and Implications for Developing Countries

In this article I examine the evidence on the effect of technology diffusion on productivity growth through imports of new intermediate capital goods. I develop an empirical model in which domestic productivity is related to the number of varieties of imported differentiated inputs that are employed domestically. Based on the hypothesis that the number of varieties from partner countries is related to imports from those countries, I estimate the relation between domestic as well as import-weighted foreign R&D and domestic productivity.

Three conclusions emerge from the analysis. First, there is evidence that countries benefit more from domestic R&D than from R&D of the average foreign country. Second, conditional on technology diffusion from domestic R&D, the import composition of a country matters, but only if it is strongly biased toward or away from technological leaders. Third, differences in technology inflows related to the patterns of imports explain about 20 percent of the total variation in countries' productivity growth rates.

What are the implications of this analysis for developing countries? The results suggesting that domestic R&D has a larger influence on productivity than

R&D investments in the average country abroad must be qualified for developing countries, many of which spend only a fraction of their total spending on technology on formal R&D. It is likely that the contribution of foreign sources of technology is larger than that of domestic sources for many developing countries.

To confirm this conjecture requires high-quality industry-level measures of productivity and technological efforts in developing countries, which are often difficult to obtain. The conjecture that the relative contribution of foreign sources of technology is higher the smaller is the country's relative contribution to the world's pool of technological knowledge seems to be confirmed, however, by results in Keller (2000). There, I estimate that in nine small OECD countries, the R&D of the G-5 countries (France, Germany, Japan, the United Kingdom, and the United States) taken together leads to productivity effects that are more than twice as large as those from own-country R&D investments.

Given the greater relative importance of foreign sources of technology for a typical developing country relative to the countries in this sample, one should expect differences in overall import share and import composition to have a stronger effect on differences in productivity growth in developing countries than I have estimated here. Intermediate input imports contribute to the international diffusion of technology and hence to the transfer of technology to developing countries. Everything else equal, a higher share of trade promotes that process.

The composition of imports matters. Productivity growth in a typical developing country might not depend too much on whether 50 percent of its imports come from the United States and 30 percent from Japan, or 30 percent from the United States and 50 percent from Japan. But productivity is likely to be much lower if the country were to significantly reduce the share of its imports from both the United States and Japan while increasing its share of imports from other developing countries that are not world technology leaders. The results obtained here suggest that relative import shares help explain productivity growth even in importing OECD countries. The impact on productivity of a change in import composition is likely to be an order of magnitude larger if developing-country trade patterns shift substantially between today's technological leaders and followers.

Appendix A. Data on Import Flows

The specialized machinery trade data come from OECD (1980). Import data for the first five industries are from mid-1980. For the sixth industry I was unable to obtain data for 1980 from SITC revision 2 and therefore use 1975 data from SITC revision 1. I use these tables to derive the variable m^{v}_{wj}, the bilateral import shares of country v with countries $w \neq v$ in sector j.

Table A-1. *Food-Processing Machinery Imports (SITC 727)*
(1980 U.S.$)

	Importing country							
Exporting country	*Canada*	*France*	*Germany*	*Italy*	*Japan*	*Sweden*	*United Kingdom*	*United States*
Canada	0	938	113	7	2	26	915	4,141
France	1,398	0	7,682	7,231	1,050	837	4,631	3,960
Germany	8,513	30,099	0	18,442	11,268	11,446	30,004	36,143
Italy	4,292	22,397	10,812	0	2,403	1,461	7,634	9,431
Japan	290	38	1,832	1,709	0	156	728	8,114
Sweden	1,181	1,332	1,225	606	487	0	2,310	1,916
United Kingdom	3,655	6,274	4,638	3,226	1,679	1,800	0	8,551
United States	63,235	12,559	6,196	2,838	8,458	2,022	23,435	0

Table A-2. *Textiles and Leather Machinery Imports (SITC 724)*
(1980 U.S.$)

	Importing country							
Exporting country	*Canada*	*France*	*Germany*	*Italy*	*Japan*	*Sweden*	*United Kingdom*	*United States*
Canada	0	801	660	1,232	207	38	2,140	21,275
France	4,151	0	38,542	49,901	3,465	1,353	28,705	34,619
Germany	22,409	187,433	0	259,344	55,555	31,400	116,170	262,163
Italy	23,122	78,772	68,873	0	15,124	6,155	67,436	68,070
Japan	11,110	28,372	39,932	22,546	0	1,966	40,419	139,266
Sweden	3,558	5,145	8,530	2,181	3,864	0	9,713	29,519
United Kingdom	9,953	40,817	47,110	42,585	8,856	6,632	0	53,270
United States	143,551	27,501	33,617	21,479	14,106	5,167	49,242	0

Table A-3. *Paper and Pulp Mill Machinery Imports (SITC 725)*
(1980 U.S.$)

	Importing country							
Exporting country	*Canada*	*France*	*Germany*	*Italy*	*Japan*	*Sweden*	*United Kingdom*	*United States*
Canada	0	1,352	278	304	2,722	85	919	35,110
France	534	0	25,553	16,619	109	4,560	13,245	10,01
Germany	9,767	65,245	0	47,290	17,197	31,354	61,340	68,760
Italy	794	32,561	22,365	0	353	1,834	10,028	6,125
Japan	2,829	315	7,392	925	0	782	3,831	11,535
Sweden	5,245	6,911	18,014	4,779	1,572	0	7,263	21,098
United Kingdom	11,990	9,563	12,809	8,827	584	10,580	0	8,612
United States	88,992	8,093	19,794	4,411	11,152	7,982	18,720	0

Table A-4. *Printing and Bookbinding Machinery Imports (SITC 726)*
(1980 U.S.$)

	Importing country							
Exporting country	*Canada*	*France*	*Germany*	*Italy*	*Japan*	*Sweden*	*United Kingdom*	*United States*
Canada	0	441	272	543	153	309	2,663	8,537
France	944	0	13,589	13,497	169	6,140	31,642	8,090
Germany	18,467	133,716	0	105,198	77,149	41,834	141,982	143,425
Italy	2,320	26,061	21,148	0	10,622	2,711	20,418	51,072
Japan	6,224	4,786	5,332	3,420	0	2,227	21,027	60,713
Sweden	1,543	10,612	6,074	853	3,168	0	14,519	14,471
United Kingdom	19,206	25,519	19,636	19,271	5,219	9,126	0	49,020
United States	158,716	51,574	43,920	25,469	25,662	24,677	73,167	0

Table A-5. *Machine Tools and Metal-Working Machinery Imports (SITC 736 and 737)*
(1980 U.S.$)

	Importing country							
Exporting country	*Canada*	*France*	*Germany*	*Italy*	*Japan*	*Sweden*	*United Kingdom*	*United States*
Canada	0	826	1,904	295	1,636	295	8,508	117,564
France	9,137	0	110,469	50,354	4,034	11,574	48,758	46,705
Germany	41,546	318,019	0	223,334	87,011	129,441	288,330	345,307
Italy	11,821	138,858	154,121	0	6,504	23,166	77,445	63,596
Japan	30,259	44,462	122,266	8,507	0	28,770	122,686	617,156
Sweden	8,612	17,788	45,895	15,588	4,916	0	37,929	52,440
United Kingdom	41,064	58,034	83,115	44,457	6,081	25,671	0	169,590
United States	608,480	66,698	72,679	29,627	93,295	22,344	161,467	0

Table A-6. *Mining, Metal-Crushing, and Glass-Working Machinery Imports (SITC Revision 1, 7184 and 7185)*
(1975 U.S.$)

	Importing country							
Exporting country	*Canada*	*France*	*Germany*	*Italy*	*Japan*	*Sweden*	*United Kingdom*	*United States*
Canada	0	1,517	341	312	453	294	2,559	73,438
France	11,738	0	78,204	38,841	1,316	9,063	49,246	32,890
Germany	22,999	97,060	0	47,026	3,687	25,154	50,335	64,832
Italy	3,503	26,645	34,749	0	141	1,853	17,079	11,597
Japan	13,582	3,700	16,499	7,454	0	654	13,612	42,338
Sweden	12,421	10,708	22,294	9,466	1,739	0	19,019	10,239
United Kingdom	19,340	41,885	23,092	21,204	1,722	12,348	0	37,783
United States	644,606	75,425	46,474	22,419	37,028	18,733	104,457	0

Table A-7. *Summary Statistics on Machinery* R&D *Stocks*
(1985 U.S.$)

Country and industry	*Mean*	*Standard deviation*
Country		
Canada	92.79	34.75
France	417.40	222.53
Germany	1,403.29	1,148.55
Italy	123.03	143.10
Japan	919.80	670.24
Sweden	170.59	98.33
United Kingdom	665.71	305.32
United States	3,379.03	2,720.96
Industry		
Food	1,087.13	1,535.52
Textiles	1,363.09	2,036.03
Paper products	337.83	554.24
Printing	576.61	1,079.07
Minerals and basic metals	1,145.53	1,679.58
Metal products	960.72	1,457.14

Table A-8. *Summary Statistics on Annual Total Factor Productivity Growth, 1970–91*

Country and industry	*Mean*	*Standard deviation*
Country		
Canada	0.0119	0.0086
France	0.0256	0.0102
Germany	0.0278	0.0074
Italy	0.0207	0.0188
Japan	0.0137	0.0092
Sweden	0.0205	0.0198
United Kingdom	0.0203	0.0041
United States	0.0149	0.0088
Industry		
Food	0.0091	0.0087
Textiles	0.0217	0.0108
Paper products	0.0263	0.0101
Printing	0.0141	0.0096
Minerals and basic metals	0.0276	0.0156
Metal products	0.0178	0.0098

APPENDIX B. DATA ON RESEARCH AND DEVELOPMENT

The raw data on R&D expenditures come from OECD (1991). R&D surveys were not conducted annually in all countries included in the sample over the entire sample period, however. In the United Kingdom, for instance, surveys were conducted only every third year until well into the 1980s. In Germany R&D data are collected only every second year. This lack of annual data made it necessary to interpolate about 25 percent of all the R&D expenditure data.

The construction of the technology stock variable, n, is based on data on total business enterprise intramural expenditures on R&D for ISIC sector 382 (nonelectronic machinery), in constant 1985 U.S. dollars, with OECD purchasing power parity rates used for conversion. The OECD code for this series is BERD (see table 9B of OECD 1991). I use the perpetual inventory method to construct technology stocks, assuming that

$$n_{vt} = (1-\tilde{\delta})n_{vt-1} + \chi_{vt-1}, \forall v, t = 2, \ldots, 22$$

and

$$n_{v1} = \frac{\chi_{v1}}{g_v^n + \tilde{\delta} + 0.1}.$$

The rate of depreciation, $\tilde{\delta}$ is set at 0.05; g^n is the average annual growth rate of n over the period 1970–89 (the year endpoints for which data are available for all countries). Preliminary analysis using other values for the rate of depreciation, such as 0 or 0.1, shows that the rate of depreciation does not influence the estimation results significantly. The denominator in the calculation of n_1 is increased by 0.1 in order to obtain positive estimates of n_1 throughout.

APPENDIX C. DATA ON LABOR, PHYSICAL CAPITAL, AND GROSS PRODUCTION

The OECD (1994) STAN database is the basic source for labor, physical capital, and gross production. It provides internationally comparable data on industrial activity by sector, including data on labor input, labor compensation, investment, production, and gross production for up to 49 three-digit ISIC industries (revision 2). STAN data are OECD estimates based on data submitted by OECD member countries. The OECD has tried to ensure international comparability (see OECD 1994).

In constructing the TFP variable, F, I consider only inputs of labor and physical capital (there are no data on human capital by industry). Data on labor inputs, l, are taken directly from the STAN database (number of workers employed). This includes employees as well as the self-employed, owner proprietors, and unpaid family workers. Data on physical capital stocks are not available in that database, but data on gross fixed capital formation in current prices are. I first convert the investment flows into constant 1985 prices, using output deflators (in-

vestment good deflators were not available). The output deflators are derived from figures on value added in both current and constant 1985 prices, both of which are included in the STAN database. The capital stocks are then estimated using the perpetual inventory method. Suppressing the industry subscripts,

$$k_{vt} = (1-\hat{\delta}_v)k_{vt-1} + inv_{vt-1}, \forall v, t = 2,\ldots,22$$

and

$$k_{v1} = \frac{inv_{v1}}{g_v^{inv} + \hat{\delta}_v}, \forall v$$

where *inv* is gross fixed capital formation in constant prices (land, buildings, machinery and equipment); g^{inv} is the average annual growth rate of *inv* over the period 1970–91; and $\hat{\delta}$ is the rate of depreciation of capital. I use the following country-specific depreciation rates, taken from Jorgenson and Landau (1993): Canada, 8.51 percent; France, 17.39 percent; Germany, 17.4 percent; Italy, 11.9 percent; Japan, 6.6 percent; Sweden, 7.7 percent; the United Kingdom, 8.19 percent; and the United States, 13.31 percent. These figures, which are used throughout, are estimates for machinery in manufacturing in 1980.

According to equation 1, α_{vjt} is the share of the labor cost in production. Following the approach suggested by Hall (1990), the values of α_{vjt} are not calculated as the ratio of total labor compensation to value added (the revenue-based factor shares), both of which are included in the STAN database. Rather, using the framework of the integrated capital taxation model of King and Fullerton (see Jorgenson 1993, Fullerton and Karayannis 1993, and data provided in Jorgenson and Landau 1993), I construct cost-based factor shares that are robust in the presence of imperfect competition. The effective marginal corporate tax rate, τ, is given by the wedge between the before-tax (p_k) and after-tax rate of return (ρ)

$$\text{(C-1)} \qquad \tau = \frac{p_k - \rho}{p_k}.$$

Here, the variable of interest is p_k, the user cost of capital. It is a function of such factors as the statutory marginal tax rate on corporate income, available investment tax credits, and the rates of depreciation.

In the case of equity financing, the after-tax rate of return is

$$\text{(C-2)} \qquad \rho = \iota + \pi$$

where ι is the real interest rate and π is the rate of inflation. Jorgenson (1993) tabulates the values for the marginal effective corporate tax rate, τ. I use the so-called "fixed-r" strategy ("fixed ι" in my notation), where one gives as an input a real interest rate and deduces τ. In this case I use a value of 0.1 for the real interest rate, which, together with the actual values of π, allows me, using equations (C-1) and (C-2) to infer p_k, the user cost of capital. I use Jorgenson's values

on manufacturing (the 1980 values are used for 1970–82 in my sample, the 1985 values are used for 1983–86, and 1990 values are used for 1987–91). This clearly introduces an error. In addition, Jorgenson's values are derived from a "fixed-p" approach, as opposed to the "fixed-r" approach employed here. Moreover, the results depend on the real interest rate chosen. Finally, τ varies by asset type, and ρ is a function of the type of financing used (equity versus debt primarily). These shortcomings in the construction of the cost-based factor shares are unavoidable given the lack of more detailed data.

Fullerton and Karayannis (1993) present a sensitivity analysis in several dimensions. I have experimented myself with different values for ι and found that the basic results presented above do not depend on a particular choice for ι. The main advantage of this approach is that it uses all data on the user cost of capital compiled in Jorgenson and Landau (1993) to arrive at a productivity index that is robust to deviations from perfect competition.

To obtain robust wage shares, α, I deflate the current price of labor costs, wl, available in the STAN database (again using sectoral output deflators):

$$\alpha = \frac{wl}{wl + p_k k}.$$

Labor and capital inputs together with the factor shares allow me to construct a Thornqvist index of total inputs I_t:

$$\log\left(\frac{I_{vjt}}{I_{vjt-1}}\right) = \tfrac{1}{2}\left[\alpha_{vjt} + \alpha_{vjt-1}\right]\log\left(\frac{l_{vjt}}{l_{vjt-1}}\right) + \tfrac{1}{2}\left[\left(1-\alpha_{vjt}\right) + \left(1-\alpha_{vjt-1}\right)\right]\log\left(\frac{k_{vjt}}{k_{vjt-1}}\right).$$

This index gives a series of growth of total factor inputs. Calculating log differences of year-to-year gross real production and taking the difference between this figure and total input growth results in the TFP growth series. A value of 100 in 1970 is chosen for each of the 8×6 time series for all industries j and countries v.

Appendix D. Relationship of Monte Carlo Experiments and Average R&D Spillover Regression

Consider, for simplicity, the model above with only one regressor (industry and time subscripts are suppressed):

$$\frac{\Delta F_v}{F_v} = \alpha_0 + \beta_1 \theta_w^v(b)\frac{\Delta S_w}{S_w} + \varepsilon_v.$$

Let

$$\theta_w^v(b) = \sigma(\bar{b}) + \eta_w^v(b), \forall b$$

where $\eta_w^v(b)$ is the deviation of the trade share from its expected value (partner country by partner country) of 1/7. Then the OLS estimate of $\beta_1(b)$ equals

$$\beta_1(b) = \frac{\Sigma_v\left(\theta_w^v(b)\frac{\Delta S_w}{S_w}\frac{\Delta F_v}{F_v}\right)}{\Sigma_v\left(\theta_w^v(b)\frac{\Delta S_w}{S_w}\right)^2} = \frac{\Sigma_v\left(\sigma(\bar{b})\frac{\Delta S_w}{S_w}\frac{\Delta F_v}{F_v} + \eta_w^v(b)\frac{\Delta S_w}{S_w}\frac{\Delta F_v}{F_v}\right)}{\Sigma_v\left(\left[\sigma(\bar{b}) + \eta_w^v(b)\right]\frac{\Delta S_w}{S_w}\right)^2}, \forall b.$$

If the denominator is approximated by $\Sigma_v\left(\frac{\Delta S_w}{S_w}\right)^2\left[\sigma(\bar{b})\right]^2, \forall b, v$, this means that the average of the Monte Carlo estimates, $\beta_1(\bar{b}) = \frac{1}{B}\Sigma_b\beta_1(b)$, equals

$$\beta_1(\bar{b}) \simeq \frac{\Sigma_{b=1}^B\Sigma_v\left(\sigma(\bar{b})\frac{\Delta S_w}{S_w}\frac{\Delta F_v}{F_v} + \eta_w^v(b)\frac{\Delta S_w}{S_w}\frac{\Delta F_v}{F_v}\right)}{B\Sigma_v\left(\frac{\Delta S_w}{S_w}\right)^2\left[\sigma(\bar{b})\right]^2}.$$

The right side of the expression can be rewritten to obtain

$$\text{(D-1)} \qquad \beta_1(\bar{b}) \simeq \frac{\Sigma_v\,\sigma(\bar{b})\frac{\Delta S_w}{S_w}\frac{\Delta F_v}{F_v}}{\Sigma_v\left(\frac{\Delta S_w}{S_w}\right)^2\left[\sigma(\bar{b})\right]^2} + \frac{\Sigma_v\frac{\Delta S_w}{S_w}\frac{\Delta F_v}{F_v}\Sigma_{b=1}^B\eta_w^v(b)}{B\Sigma_v\left(\frac{\Delta S_w}{S_w}\right)^2\left[\sigma(\bar{b})\right]^2}.$$

Because $\Sigma_{b=1}^B\eta_w^v(b) = 0$, however, the second term in expression D-1 will drop out, so that $\beta_1(\bar{b})$ is approximately equal to the OLS estimate of projecting $\Delta F_v/F_v$ on $\sigma(\bar{b})\Delta S_w/S_w$. Clearly, how good the approximation above is depends on how large $\left[\Delta S_w/S_w\right]^2\left(\left[\eta_w^v(b)\right]^2 + 2\eta_w^v(b)\sigma(\bar{b})\right)$ is, or, more generally, $\lambda_w^2\left(\left[\eta_w^v(b)\right]^2 + 2\eta_w^v(b)\sigma(\bar{b})\right)$. In particular, if $\lambda_w = \log S_w$, then the average Monte Carlo estimate will differ more from the average spillover regression than if $\lambda_w = \Delta S_w/S_w$, the case presented in table 4.

APPENDIX E. SENSITIVITY ANALYSIS

The sensitivity of the results can be examined by considering a number of alternative specifications for both the productivity level and the growth regressions. As noted above, in principle, including a fixed effect for each industry allows consistent parameters to be estimated with OLS if the error is of the form $\varepsilon_{vjt} = u_{vj} + \eta_t$, because the correlation between error and regressor due to u_{vj} will be subsumed into the fixed effects. Including a separate fixed effect for every industry leads to the following specifications, analogous to equations 8 and 9:

$$\log F_{vjt} = \delta_{vj} d_{vj} + \sum_{c \in G7S} \beta_c (m_{cj}^v \log S_{cjt}) + \varepsilon_{vjt}, \forall v, j, t$$

and

$$\log F_{vjt} = \delta_{vj} d_{vj} + \sum_{c \in G7S} \beta_c (m_{vj} m_{cj}^v \log S_{cjt}) + \varepsilon_{vjt}, \forall v, j, t$$

where d_{vj} is dummy variable that equals 1 for the country-industry combination vj and zero otherwise. Including a separate fixed effect for every industry raises the number of fixed effects from 14 to 48 (6×8). The inclusion of more fixed effects raises the R^2 for both specifications, from 0.472 to 0.755 for the NIS and from 0.357 to 0.746 for the IS. In both specifications all of the estimated parameters β_c remain significantly different from zero at the 1 percent level, with the values ranging from 4.1 percent (for the United States in the NIS) to 61.5 percent (for France in the IS).

Inclusion of a trend (denoted *year*) yields

$$\text{(E-1)} \qquad \log F_{vjt} = \alpha year_t + \delta_{vj} d_{vj} + \sum_{c \in G7S} \beta_c (m_{cj}^v \log S_{cjt}) + \varepsilon_{vjt}, \forall v, j, t$$

and

$$\text{(E-2)} \qquad \log F_{vjt} = \alpha year_t + \delta_{vj} d_{vj} + \sum_{c \in G7S} \beta_c (m_{vj} m_{cj}^v \log S_{cjt}) + \varepsilon_{vjt}, \forall v, j, t$$

The trend increases the R^2 slightly (from 0.755 to 0.757 in the NIS and from 0.746 to 0.753 in the IS). It also lowers the estimated parameters β_c in both specifications—just as one would expect if there are common trends in levels. The new estimates range from a low of 2.4 percent (for the United States in the NIS) to a high of 44.1 percent (for France in the NIS). On average, the estimates fall by about 15–20 percent. However, 13 of 16 estimates from the specifications given by equations E-1 and E-2 remain significantly positive at the 5 percent level; the highest p-value is 21.1 percent.

In the growth specifications a major concern is whether all industries share a common growth rate. To test this, I run equations 16 and 17 including industry fixed effects to get:

$$\text{(E-3)} \qquad \frac{\Delta F_{vjt}}{F_{vjt}} = \delta_j d_j + \sum_{c \in G7S} \beta_c \left(m_{cj}^{v} \frac{\Delta S_{cjt}}{S_{cjt}} \right) + \varepsilon_{vjt}, \forall v, j, t$$

and

$$\text{(E-4)} \qquad \frac{\Delta F_{vjt}}{F_{vjt}} = \delta_j d_j + \sum_{c \in G7S} \beta_c \left(m_{vj} m_{cj}^{v} \frac{\Delta S_{cjt}}{S_{cjt}} \right) + \varepsilon_{vjt}, \forall v, j, t.$$

For specification E-3, one out of six industry fixed effects is estimated to be significantly different from zero at the 5 percent level. For specification E-4, none of the industry fixed effects is significant. The estimated parameters β_c are affected only slightly.

Including country fixed effects in addition to the industry fixed effects leads to the new specifications:

$$\text{(E-5)} \qquad \frac{\Delta F_{vjt}}{F_{vjt}} = \mu_v d_v + \delta_j d_j + \sum_{c \in G7S} \beta_c \left(m_{cj}^{v} \frac{\Delta S_{cjt}}{S_{cjt}} \right) + \varepsilon_{vjt}, \forall v, j, t$$

and

$$\text{(E-6)} \qquad \frac{\Delta F_{vjt}}{F_{vjt}} = \mu_v d_v + \delta_j d_j + \sum_{c \in G7S} \beta_c \left(m_{vj} m_{cj}^{v} \frac{\Delta S_{cjt}}{S_{cjt}} \right) + \varepsilon_{vjt}, \forall v, j, t.$$

In specification E-5, 4 out of 14 fixed effects differ significantly from zero. In specification E-6, 3 out of 14 fixed effects are significant. These results suggest that the evidence that growth rates differ across industries is not very strong. Even when one includes country-by-industry fixed effects (that is, $6 \times 8 = 48$ fixed effects), only about 30 percent of these values are estimated to be significantly different from zero.

Including more fixed effects does reduce the number of R&D parameters β_c that are estimated to differ significantly from zero. In the growth specification with overall import share, for example, when country-by-industry fixed effects are included, only the R&D stocks of Canada, Germany, the United Kingdom, and the United States are estimated to have a significantly positive effect on productivity (at the 10 percent level). This is to be expected in a cross-industry, cross-country TFP growth regression that does not exploit any between-industry variation. Overall, this analysis suggests that the results presented in the text are fairly robust.

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Aghion, Philippe, and Peter Howitt. 1998. *Endogenous Growth Theory*. Cambridge, Mass: MIT Press.

———. 1992. "A Model of Growth through Creative Destruction." *Econometrica* 60(2):323–51.

Barro, Robert, and Xavier Sala-i-Martin. 1995. *Economic Growth*. New York: McGraw-Hill.

Bernstein, Jeffrey, and Pierre Mohnen. 1998. "International R&D Spillovers between U.S. and Japanese R&D Intensive Sectors." *Journal of International Economics* 44(2):315–38.

Branstetter, Lee. Forthcoming. "Are Knowledge Spillovers International or Intranational in Scope? Microeconometric Evidence from the U.S. and Japan," *Journal of International Economics*.

Coe, David, and Elhanan Helpman. 1995. "International R&D Spillovers." *European Economic Review* 39(5):859–87.

Coe, David, Elhanan Helpman, and Alexander Hoffmaister. 1997. "North-South R&D Spillovers." *Economic Journal* 107(440):134–49.

Eaton, Jonathan, and Samuel Kortum. Forthcoming. "International Patenting and Technology Diffusion: Theory and Measurement." *International Economic Review*.

Edwards, Sebastian. 1993. "Openness, Trade Liberalization, and Growth in Developing Countries." *Journal of Economic Literature* 31(3):1358–93.

Ethier, Wilfred. 1982. "National and International Returns to Scale in the Modern Theory of International Trade." *American Economic Review* 72(3):389–405.

Evenett, Simon, and Wolfgang Keller. 1998. "On Theories Explaining the Success of the Gravity Equation." NBER Working Paper No. 6529. National Bureau of Economic Research, Cambridge, Mass.

Evenson, Robert. 1995. "Industrial Productivity Growth Linkages Between OECD Countries: 1970–1990." Department of Economics, Yale University, New Haven, Conn.

Feenstra, Robert C. 1996. "Trade and Uneven Growth." *Journal of Development Economics* 49(1):229–56.

Fullerton, Don, and Marios Karayannis. 1993. "United States." In Dale W. Jorgenson and Ralph Landau, eds., *Tax Reform and the Cost of Capital*. Washington, D.C.: Brookings Institution.

Griliches, Zvi. 1979. "Issues in Assessing the Contribution of Research and Development to Productivity Growth." *Bell Journal of Economics* 10(1):92–116.

———. 1995. "R&D and Productivity: Econometric Results and Measurement Issues." In Paul Stoneman, ed., *Handbook of the Economics of Innovation and Technological Change*. Oxford: Blackwell.

Griliches, Zvi, and Jerry Hausman. 1986. "Errors in Variables in Panel Data." *Journal of Econometrics* 31(1):93–118.

Grossman, Gene M., and Elhanan Helpman. 1990. "Comparative Advantage and Long-run Growth." *American Economic Review* 80(4):796–815.

———. 1991. *Innovation and Growth in the Global Economy*. Cambridge, Mass: MIT Press.

Hall, Robert E. 1990. "Invariance Properties of Solow's Productivity Residual." In Peter Diamond, ed., *Growth/Productivity/Employment*. Cambridge, Mass: MIT Press.

Helpman, Elhanan. 1997. "R&D and Productivity: The International Connection." NBER Working Paper No. 6101. National Bureau of Economic Research, Cambridge, Mass.

Jorgenson, Dale W. 1993. "Introduction and Summary." In Dale W. Jorgenson and Ralph Landau, eds., *Tax Reform and the Cost of Capital.* Washington, D.C.; Brookings Institution.

Jorgenson, Dale W., and Ralph Landau. 1993. "Appendix." In Dale W. Jorgenson and Ralph Landau, eds., *Tax Reform and the Cost of Capital.* Washington, D.C: Brookings Institution.

Keller, Wolfgang. 1996. "Absorptive Capacity: On the Creation and Acquisition of Technology in Development." *Journal of Development Economics* 49(1):199–227.

———. 1997. "How Trade Patterns and Technology Flows Affect Productivity Growth." Policy Research Working Paper No. 1831. Development Research Group, World Bank, Washington, D.C.

———. 1998a. "Are International R&D Spillovers Trade-Related? Analyzing Spillovers among Randomly Matched Trade Partners." *European Economic Review* 42(8):1469–81.

———. 1998b. "Trade and the Transmission of Technology." Department of Economics, University of Wisconsin, Madison. Processed.

Keller, Wolfgang. 2000. "Geographic Localization in International Technology Diffusion." NBER Working Paper 7509, National Bureau of Economic Research, Cambridge, Mass.

Klenow, Peter, and Andres Rodriguez-Clare. 1996. "Quantifying Variety Gains from Trade Liberalization." Graduate School of Business, University of Chicago.

Lichtenberg, Frank. 1993. "R&D Investment and International Productivity Differences." In Horst Siebert, ed., *Economic Growth in the World Economy.* Tuebingen, Germany: J. C. B. Mohr.

Lichtenberg, Frank, and Bruno van Pottelsberghe de la Potterie. 1998a. "Does Foreign Direct Investment Transfer Technology across Borders?" Columbia University, New York City. Processed.

———. 1998b. "International R&D Spillovers: A Comment." *European Economic Review* 42(8):1483–91.

OECD (Organisation for European Co-operation and Development). 1980. *Trade by Commodities.* Paris.

OECD. 1991. *Basic Science and Technology Statistics.* Paris.

———. 1994. STAN *Industrial Database Analysis* (Structural Analysis International Database). Paris.

Park, Walter. 1995. "International R&D Spillovers and OECD Economic Growth." *Economic Inquiry* 33(4):571–91.

Rivera-Batiz, Luis, and Paul Romer. 1991. "Economic Integration and Endogenous Growth." *Quarterly Journal of Economics* 106(2):531–55.

Romer, Paul. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98(5):S71–S102.

Wang, J., and B. Xu. 1996. "International R&D Spillovers: The Role of Capital Goods Trade and Foreign Direct Investment." Department of Economics Working Paper 96-97-5, University of Florida, Gainesville.

World Bank. 1991. *World Development Report 1991: The Challenge of Development.* New York: Oxford University Press.

———. 1998. *World Development Report 1998/99: Knowledge for Development.* New York: Oxford University Press.

THE WORLD BANK ECONOMIC REVIEW, VOL. 14, NO. 1: 49–64

Foreign Investment and Productivity Growth in Czech Enterprises

Simeon Djankov and Bernard Hoekman

This article uses firm-level data for the Czech Republic to show that during 1992–96 foreign investment had the predicted positive impact on total factor productivity growth of recipient firms. This result is robust to corrections for the sample bias that arises because foreign companies tend to invest in firms whose initial productivity is above average. Together, joint ventures and foreign direct investment appear to have a negative spillover effect on firms that do not have foreign partnerships. However, with foreign direct investment alone, the magnitude of the spillover becomes much smaller and loses significance. This result, in conjunction with the fact that joint ventures and foreign direct investment account for a significant share of total output in many industries, suggests that further research is required to determine the extent of knowledge diffusion from firms that have foreign links to those that do not.

There is a rich case-study literature documenting how firms and industries adopt new technology and knowledge. It points out that imports and openness to trade are vital to learning, which is achieved through reverse-engineering, direct inputs into production, and communication with foreign partners (suppliers and buyers). A number of recent studies that use aggregate data conclude that trading with countries that are relatively intensive in research and development (R&D) leads to higher productivity growth in domestic industry (Coe and Helpman 1995 and Coe, Helpman, and Hoffmaister 1997). These findings are consistent with the endogenous growth literature, although they do not reveal much about *how* technologies are transferred from one country to another.

The microeconomic literature emphasizes three channels through which technologies are transferred internationally: imports of new capital and differentiated intermediate goods (Feenstra, Markusen, and Zeile 1992 and Grossman and Helpman 1995), learning by exporting (Clerides, Lach, and Tybout 1998), and foreign investment (Blomström and Kokko 1997). Particular attention has

Simeon Djankov is with the Financial Sector Practice Department at the World Bank, and Bernard Hoekman is with the World Bank Institute and the Centre for Economic Policy Research. Their e-mail addresses are *sdjankov@worldbank.org* and *bhoekman@worldbank.org*. An earlier version of this article was presented at the conference Trade and Technology Diffusion: The Evidence with Implications for Developing Countries, sponsored by the Fondazione Eni Enrico Mattei and the International Trade Division of the World Bank and held in Milan on April 18–19, 1997. The authors are grateful to Magnus Blomström, Caroline Freund, Ann Harrison, Roberto Rocha, Jim Tybout, and three anonymous referees for helpful comments and suggestions.

centered on the role of foreign investment as a channel for the transfer of knowledge and on the spillover of this knowledge to other firms in the economy.[1] Foreign investment should be associated with the transfer of knowledge because, by definition, it is driven by intangible assets owned by the parent firm (Markusen 1995). The conventional wisdom holds that foreign investment is a major channel for technology transfer to developing countries. Pack and Saggi (1997) note that transactions in royalties and license fees between parent firms and subsidiaries account for more than 80 percent of global flows of foreign investment.

What matters for economic growth are the spillovers to other firms within and across industries. Evidence on this issue is much less robust. Case studies have argued that positive spillovers are significant. They also have documented the importance of local skills and in-house technological capacity for adapting and using techniques developed elsewhere (Lall 1992 and Evenson and Westphal 1995). However, recent microeconometric studies using firm-level panel data have reached more ambiguous conclusions. Some analysts have found a statistically significant negative relationship between the value of foreign investment in an industry or economy and the productivity of domestic firms (see, for example, Harrison 1996 and Haddad and Harrison 1993).

This article investigates how foreign investment affected the productivity of firms in the Czech Republic during the initial post-reform period (1992–96). We distinguish between Czech firms that established partnerships with foreign firms—either through joint ventures or through direct sales of a majority equity stake—and those that did not and ask whether the total factor productivity (TFP) growth rates of these groups differed.[2]

I. Channels of Technology Transfer

Although there is little doubt that technologies make their way across international borders, the mechanisms through which this occurs are poorly understood. Aside from case studies, most of the empirical evidence is based on aggregate data or cross-sectional surveys and is subject to multiple interpretations. Technologies may be transferred through several channels. New technologies may be embodied in new varieties of differentiated products or capital goods and equipment. They may be transferred through imports or through arm's-length trade in intellectual property, such as licensing contracts. Firms may learn about new technologies by exporting to knowledgeable buyers who share product designs

1. A separate but related literature on technology diffusion has focused largely on two issues: the determinants of the number of firms or the proportion of industry output produced by a new technology (aggregate diffusion) and the determinants of the time at which a firm adopts a new technology relative to other firms (the so-called duration models). See, for example, Ray (1964) and Karshenas and Stoneman (1994). We cannot analyze the types of questions asked in the diffusion literature because we cannot identify specific technologies in our data set.

2. TFP is an indirect measure of technology transfer. Data constraints prevent us from using more direct measures, such as investment in R&D or the turnover of managers and highly skilled labor.

and production techniques. Technologies also may be transferred in the context of formal cooperative arrangements between foreign and local firms, such as foreign direct investment (FDI) (acquisition) or project-specific joint ventures.[3] In all of these cases absorbing and adapting new technologies require workers who have appropriate training and expertise. The absence of such capacity is often held to explain why TFP frequently is lower in developing-country firms than in industrial-country firms, even if both use identical equipment (Pack 1987).

It is helpful to differentiate between technology transfers that are made in the context of formal cooperative arrangements between a foreign and a domestic firm and those that occur at arm's length. The latter, which include arm's-length trade in machinery and components and direct purchases of knowledge (payment for patents, blueprints, and so on), can be a major avenue of technology transfer. However, not all technologies are available at arm's length. Some may be obtainable only through formal cooperation—either majority ownership (acquisition) or project-specific joint ventures.[4] In theory, firms will be adverse to unbundling and selling knowledge or products if there are important incentives for internalization—in this case FDI may be the preferred channel for acquiring knowledge (Markusen 1995, 1998).

Foreign investment is likely to be associated with the transfer of both hard (machinery, blueprints) and soft (management, information) technologies. It has two dimensions: generic knowledge, such as management skills and quality systems, and specific knowledge, which cannot be obtained at arm's length because of weaknesses in the receiving country's policy environment (such as poor enforcement of intellectual property rights) or because of incentives for internalization.[5] As for generic knowledge, foreign partners may reduce the cost of learning and upgrading by helping to identify and implement systems to ensure that the product meets technical specifications, is delivered on time, and so on. Our interviews with managers of enterprises that have foreign partnerships suggest that all of these dimensions are prevalent in the Czech Republic. Still more important is access to information specific to the parent firm, as well as production and distribution networks.

An important question is whether and the extent to which knowledge that multinationals transfer to affiliates diffuses to other firms in the industry.[6] Theoretical models of foreign investment suggest that there should be a positive rela-

3. See for example, Helleiner (1973) and Keesing and Lall (1992) on subcontracting; Feenstra, Markusen, and Zeile (1992) on imports of inputs; Blomström and Kokko (1997) for a recent survey of the literature on FDI; and Pack and Saggi (1997) for a general survey of the literature on technology transfer.

4. Notions of arm's-length exchange used in the literature vary. For example, Pack and Saggi (1997) distinguish between intrafirm exchange (FDI) and contractual exchange (licensing, joint ventures, turnkey projects). They call contractual exchange arm's-length arrangements.

5. See Smarzynska (1998) for a recent analysis of the relationship between intellectual property protection and FDI in transition economies.

6. Equally important may be spillovers across industries. We do not explore this issue here, although it may be important in the context of transition.

tionship between FDI and diffusion. Knowledge will move from firm to firm through demonstration effects, labor turnover, or reverse-engineering. Das (1987) models a foreign subsidiary as the price leader and domestic firms as the competitive fringe. If the learning of domestic firms is proportional to the output of the multinational firm—that is, the larger the multinational is relative to the domestic industry, then the easier learning is—the multinational firm has an incentive to transfer technologies to its subsidiary since more advanced technologies raise profits. The greater output of the subsidiary then induces local firms to learn and adopt the foreign technologies at a faster rate. Wang and Blomström (1992) use a similar setup, but endogenize both the amount of technology transferred from the parent company to the subsidiary and the domestic firms' investment in learning activities. Foreign firms again transfer technologies at a higher rate if domestic firms invest more in learning activities. Blomström, Kokko, and Zejan (1994) find some empirical support for this prediction.

The empirical evidence on spillovers from foreign-owned affiliates to indigenous firms is mixed (Blomström and Kokko 1997). An extensive case-study literature seeks to determine the size of spillovers from R&D, if any. Much of this literature focuses on industrial countries.[7] The studies on developing countries reveal that the magnitude of potential knowledge spillovers depends on the technological capabilities of indigenous firms that would enable them to assimilate knowledge (Pack and Westphal 1986). A unique feature of many transition economies compared to most developing countries is that their technological ability is substantially greater. In principle, this should facilitate the adoption of new technologies and allow rapid convergence toward best practice.

Much of the econometric literature has focused on productivity measures as proxies for measures of technology diffusion. Early studies using industry-level data, such as Blomström and Persson (1983), find that foreign presence in an industry, measured by the foreign share of industry employment, positively influences domestic labor productivity. More recent studies using firm-level data are less supportive of the existence of spillovers. Aitken, Hanson, and Harrison (1997) and Haddad and Harrison (1993) find that foreign investment has a negative effect on the performance of domestically owned firms. Harrison (1996) suggests that in imperfectly competitive markets entry by foreign investors implies that domestic incumbents lose market share, impeding their ability to attain scale economies. The result showing negative spillovers contrasts with the findings of the case-study literature and may to some extent reflect the omission of important variables, such as the level of R&D spending, expenditures on training, and the percentage of employees with technical degrees (engineers, scientists).[8]

7. See Griliches (1992) for a survey of the literature on R&D spillovers and Nelson and Wolff (1997) for a recent contribution to this literature.

8. The literature on acquiring and adopting technologies in developing countries is substantial. See, for example, Evenson and Westphal (1995), Lall (1987, 1992), and Pack and Westphal (1986). Westphal, Rhee, and Pursell (1981) discuss the case of the Republic of Korea in some depth.

In this article we estimate production functions using TFP as a proxy for technology transfer. By relying on TFP as the dependent variable, we assume that the adoption of new technologies will, with some lag, improve productivity. A serious problem with this assumption is that, as the case-study literature has documented, such improvements depend on the technological abilities of domestic firms. Nelson and Pack (1998) demonstrate that the production function methodology can underestimate or ignore the use of improved technologies at the level of the firm and thus affect estimates of TFP growth. Differences in technological capacity across firms in an industry may be an important determinant of TFP, but we do not have this information—data on variables relevant to technology, such as R&D expenditures or the composition of the workforce, are not available at the level of the firm. However, the Czech Republic is not a developing country—it has a long-standing industrial base and is well endowed with engineering and scientific human capital. For the economy as whole, therefore, the capacity to upgrade productive efficiency rapidly by adopting best-practice techniques (both hard and soft) should be considerable.

II. A Profile of Czech Firms

We compiled information on Czech enterprises for 1992–96 from surveys using a questionnaire that we prepared and a database developed by the Czech Statistical Office containing financial and ownership information. We defined financial variables using international accounting standards from the onset of the survey in 1992. The database comprises 513 firms quoted on the Prague stock exchange whose shares traded at least four times in a given year (this restriction excludes smaller firms from the sample) and that reported the financial information required. Of the sample firms, 340 did not establish joint ventures or attract FDI, 91 concluded joint ventures with foreign companies, and 82 attracted majority foreign equity investment. Thus 34 percent of the sample (173 firms) had a foreign link—either a joint venture or FDI—with relatively uniform distribution across sectors (table 1). There is a selection bias in the data, as the sample does not cover all listed firms with foreign ownership or partnerships. Moreover, privately held firms are not included in the sample. For example, the largest foreign acquisition in the Czech Republic to date—the takeover of Skoda by Volkswagen—is not publicly traded.

To determine whether or not a firm had a foreign partnership or foreign ownership, we chose as our criterion that at least 20 percent of the equity had to be owned by a single foreign entity or the firm had to have established one or more joint ventures with a foreign partner. Because minority shareholders have little protection under Czech law, equity investors have an incentive to take a majority stake. Most firms with foreign equity ownership in the sample are majority foreign-owned. Although the share of firms with foreign links appears to be high, it is representative of Czech industry more generally. Aggregate statistics using a criterion of 5 percent or more foreign equity ownership reveal that during 1994–

Table 1. *Descriptive Statistics of Sample Firms, 1992–96*

Sector	*Total in sample*	*No foreign partner*	*Foreign partner (foreign direct investment or joint venture)*
Mining	11	8	3
Construction	82	55	27
Food and beverage	54	36	18
Textiles and apparel	39	28	11
Furniture and other wood products	11	5	6
Pulp and paper	14	10	4
Printing and publishing	13	6	7
Chemicals	30	18	12
Shoes and leather products	6	5	1
Nonmetallic mineral products	21	16	5
Basic metals	13	9	4
Fabricated metal products	24	12	12
Electric and electronics	82	54	28
Transport equipment	12	5	7
Other manufacturing	10	6	4
Retail services	15	11	4
Financial services	76	56	20
Number of observations	513	340	173
Share in total (percent)	100.0	66.3	33.7

Source: Authors' survey.

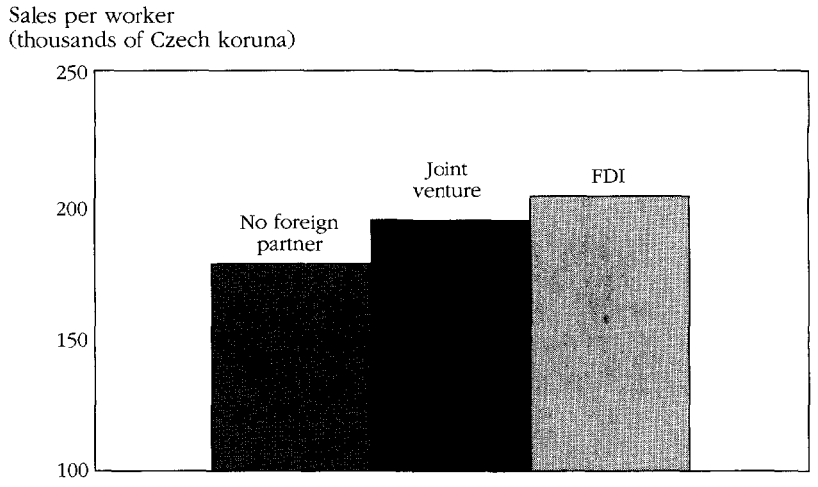
97, 42 percent of all manufacturing firms with more than 10 employees were involved in some kind of foreign partnership (Czech Statistical Office 1998).

Firms with foreign partnerships tend to be significantly larger than firms that remain independent: the median number of employees is 689 in FDI firms, 578 in firms with joint ventures, and 352 in firms without foreign links. Foreign affiliates or joint ventures also have higher initial labor productivity, measured as sales per worker in 1991 (figure 1). This suggests that foreign investors are attracted to firms with above-average performance and size.

Firms with FDI also have the highest average TFP growth of the three groups, followed by firms with joint ventures and then domestic enterprises (figure 2). This ordering may reflect the fact that the initial productivity of firms that attract foreign investment is better than average, implying that foreign investors choose the best firms as partners. In our statistical analysis we therefore correct for the possibility of selection bias. TFP growth rates are highest in earlier years and taper off toward the end of the sample period, reflecting a marked deterioration in macroeconomic conditions in 1996, a common effect for all firms. TFP growth rates initially diverge substantially; growth rates rise in firms with foreign investment and fall in others. Thereafter, some convergence occurs, suggesting that spillover effects may be in play toward the end of the period.

Our questionnaires reveal that both joint ventures and FDI are associated with technology transfers. A questionnaire sent to the sample firms in early 1997 in-

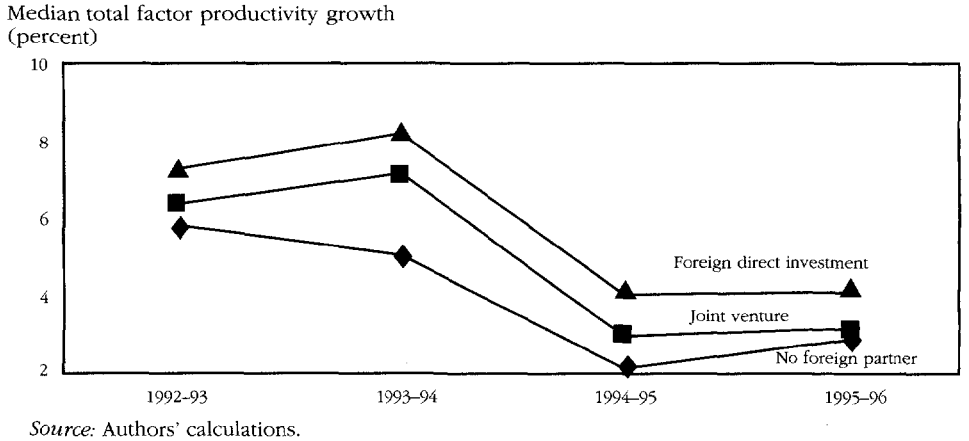
Figure 1. *Initial Labor Productivity, 1991*



Source: Authors' survey.

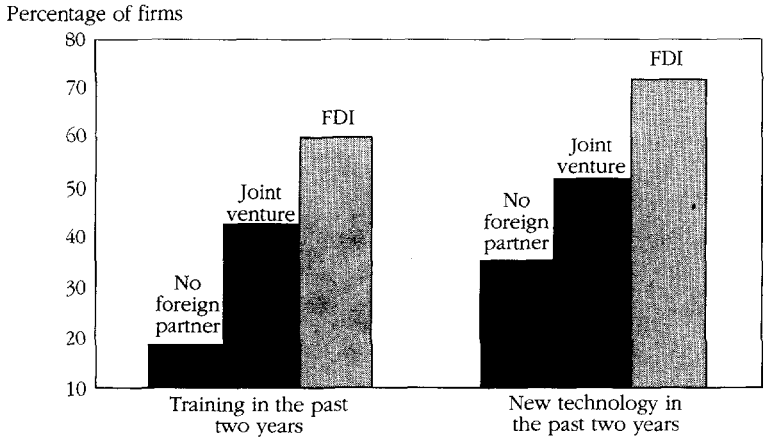
cluded two questions related to training and acquisition of new technologies (figure 3). Managerial responses clearly reveal what appears to be a significant difference between firms with and without foreign partnerships. The questionnaire first asked managers whether their workers had undergone any training in the past two years. Managers were given discrete choices: yes or no. In firms without foreign partners only 18 percent replied positively, while 42 and 62 percent of managers whose firms were involved in joint ventures or FDI, respectively, answered positively. The second question asked whether the firm had obtained new technologies (machinery, equipment) or related knowledge in the previous two years. Again, the response was similar. In more than 70 percent of the FDI firms

Figure 2. *Total Factor Productivity Growth, 1992–96*



Source: Authors' calculations.

Figure 3. *Training and Acquisition of New Technologies, 1997*



Source: Authors' survey.

and 50 percent of the joint ventures the partner had acquired some kind of new technology, as opposed to only 35 percent of firms without foreign links. The relative difference between the two sets of firms is greater for the training variable (software) than for the technology variable (hardware).

III. The Estimation Procedure

We estimate production functions for the firms included in the sample. Each firm i has a production function for gross output:

$$Y_i = F^i(K_i, L_i, M_i) \tag{1}$$

where Y is gross output, and K, L, and M are inputs of capital, labor, and materials. The firm's production function F is homogeneous of degree g $(g \neq 1)$ in K, L, and M. Firms are assumed to be price takers on factor markets, but they may have market power in output markets. The assumption that firms are price takers is reasonable since most wages were set centrally during the sample period, and most materials were bought abroad at world market prices.

The production function in equation 1 implies the following relation between marginal physical products and outputs:[9]

$$F^i_L L_i + F^i_M M_i + F^i_K K_i = g_i Y_i \tag{2}$$

where F^i_J is the marginal product of input J. The optimal choice of inputs by a firm with some monopoly power implies:

$$P_i F^i_J = \mu_i P_{Ji} \tag{3}$$

9. We are grateful to a referee for suggesting the specific formulation used below.

where P_{Ji} is the price of factor J, P_i is the price of the firm's output, and μ_i is the markup of price over marginal cost: $\mu_i = P_i/MC_i$, where MC_i is marginal cost. Combining equations 2 and 3, we obtain:

$$s_{Li} + s_{Mi} + s_{Ki} = g_i/\mu_i \tag{4}$$

where $s_{Ji} = P_{ji}J_i/P_iY_i$ are expenditures on each factor J_i relative to total enterprise revenues. Since firms do not necessarily produce under constant returns to scale, the sum of these shares is not always unity. Using equation 4, the revenue share of capital can be defined as:

$$\hat{s}_{Ki} = 1 - s_{Li} - s_{Mi} = s_{ki} + (1 - g_i/\mu_i). \tag{5}$$

The productivity equation can then be derived from equation 1 as

$$dy_i = \mu_i(s_{Li}dl_i + \hat{s}_{Ki}dk_i + s_{Mi}dm) + \mu_i(s_{Ki} - \hat{s}_{Ki})dk_i + \frac{F_T^iT_i}{F^i}dt_i \tag{6}$$

where dy_i is output growth and $(F_T^iT_i/F^i)dt_i$ measures the technology change or TFP growth not accounted for by the increase in input use. The second term on the right side can be simplified to $(g_i - \mu_i)dk_i$ using equation 5.

We estimate equation 6 in log-differences, using actual enterprise-level data to construct the first right-side term. There are two terms to estimate for each industry, g_i and μ_i, the scale and markup parameters, in addition to the TFP parameter for each enterprise. We use the reported book value of fixed assets to construct the share of capital revenue.

To correct for the likelihood that foreign investment choices are not randomly distributed, we use the generalized Heckman two-step procedure for correcting sample selection bias as developed by Amemiya (1984). This procedure involves separately estimating the foreign investment decision and the firm's subsequent productivity growth. The first step uses a probit model to determine the probability of foreign investment based on initial efficiency (proxied by the share of variable costs in total revenue), firm size, and type of industry. The second step involves estimating productivity using only observations on firms with foreign links. Because this would generate omitted variable bias, the Amemiya procedure provides a specification of the omitted variable that can be used in the full sample to alleviate sample selection. This additional variable estimated in the first step is then included as a regressor in the second step.

Since the primary focus of this article is to test for an association between productivity growth and foreign investment, we augment equation 6 by including dummies for firms with foreign partners as additional factors of production. The dummies (FDI, JV) take the value of 1 if a firm had either FDI or a joint venture in the preceding year and 0 otherwise. This approach is similar to the empirical design that Harrison (1994) uses. We also have to control for the ef-

fects of other changes in the economic environment, but we do not have good proxies for these changes, nor can we account for each of them individually. Instead, we include annual dummies in the estimating equation. These pick up the net effects of changes in the aggregate economy.

We also are interested in determining whether there have been any spillover effects from foreign investment. To do so, we run equation 6 on domestic firms only and include as an additional independent variable (called SPILLOVER) the ratio of the assets of firms with FDI or joint ventures to the assets of all firms in each sector. If foreign participation has beneficial spillover effects, we would expect the coefficient to be positive. We also run an alternative specification, grouping joint ventures together with local firms. We then estimate the spillover effects that FDI firms have on the larger group.

Because of the probable correlation between productivity and the independent variables, ordinary least squares (OLS) may give biased and inconsistent estimates. This simultaneity problem is endemic to the empirical literature on measuring productivity. We address this issue first by using *F*-tests to reveal whether or not OLS is appropriate. Then, if OLS is inappropriate, we use the Hausman specification test to choose between random- or fixed-effects frameworks.

These two tests suggest that a random-effects model is most appropriate. A fixed-effects estimation assumes that firm productivity growth is constant over time. The problem with this assumption is that we want to examine changes in productivity arising from increased competition. The random-effects model avoids this assumption, but it assumes that productivity shocks at the firm level are uncorrelated over time. This restriction may not be reasonable if there is convergence or divergence in corporate performance. Estimates for the major coefficients or variables of interest are reported in table 2.

IV. Results

We estimate equation 6 using both an OLS and a random-effects specification (table 3). The estimated coefficient on the dummy for FDI is positive and statistically significant for both specifications, suggesting that, as predicted, foreign investment involves an additional transfer of technology. The dummy for joint ventures also has a positive sign, but it is slightly smaller in magnitude and is not statistically significant.

We consider the possibility that foreign investment will have a positive spillover effect by including the share of assets of firms with foreign partners in total assets (lagged one year) as a separate regressor. This is a continuous, not a categorical, variable. This approach assumes that spillovers are sector-specific and therefore ignores possible spillovers between industries. Contrary to what is predicted, spillovers are negative: greater foreign participation in an industry has a statistically significant negative effect on the performance of other firms (table 4). Each 10 percent increase in the share of foreign assets is associated with a 1.7 percent fall in sales growth of domestic firms.

Table 2. *Revenue Shares of Inputs, Markup, and Scale Estimates, 1992–96*

Sector	*Revenue share of* *Materials*	*Labor*	*Capital*	*Markup*	*Scale estimates*	*Share of foreign assets*	*Share of FDI assets*
Mining	0.538	0.215	0.246	1.246	1.200	0.398	0.124
Construction	0.720	0.169	0.111	1.137	1.088	0.432	0.325
Food and beverage	0.629	0.206	0.165	1.388	1.264	0.635	0.311
Textiles and apparel	0.677	0.180	0.142	1.284	1.132	0.294	0.182
Furniture and other wood products	0.743	0.145	0.110	1.152	1.001	0.542	0.261
Pulp and paper	0.791	0.129	0.079	1.211	1.113	0.715	0.521
Printing and publishing	0.730	0.136	0.133	0.889	0.992	0.885	0.605
Chemicals	0.757	0.151	0.091	1.201	1.163	0.547	0.281
Shoes and leather products	0.612	0.224	0.162	1.182	1.119	0.128	0.000
Nonmetallic mineral products	0.615	0.191	0.193	0.958	0.996	0.408	0.241
Basic metals	0.702	0.155	0.142	1.211	0.880	0.367	0.134
Fabricated metal products	0.733	0.121	0.145	1.192	1.100	0.785	0.191
Electric and electronics	0.657	0.191	0.151	1.201	1.039	0.356	0.110
Transport equipment	0.687	0.117	0.195	1.272	1.070	0.428	0.127
Other manufacturing	0.594	0.171	0.233	n.a.	n.a.	0.524	0.229
Retail services	0.257	0.453	0.289	1.352	1.198	0.402	0.221
Financial services	0.190	0.609	0.200	1.079	1.324	0.368	0.141
Average	0.625	0.209	0.164	1.184	1.104	0.483	0.191

n.a. Not applicable.
Source: Authors' calculations.

Table 3. *Panel Regression Estimates (Full Sample)*

Dependent variable: Growth in sales	*Ordinary least squares estimation*	*Random-effects estimation*
Amemiya selection bias correction variable	Yes	Yes
Sector-specific returns to scale and markups	Yes	Yes
Foreign direct investment dummy	0.015**	0.015*
	(2.011)	(1.937)
Joint venture dummy	0.011	0.010
	(1.372)	(1.286)
Dummy for 1994	–0.012*	–0.011
	(–1.873)	(–1.672)
Dummy for 1995	–0.052**	–0.052**
	(–7.034)	(–6.942)
Dummy for 1996	–0.054**	–0.053**
	(–7.062)	(–7.534)
Number of observations	513	513
F-test ($A, B = A_i, B$)	0.89	
Hausman test (random versus fixed effects)[a]		25.66 [30.19]
Adjusted R^2	0.894	0.861

* Significant at the 10 percent level.
** Significant at the 5 percent level.
Note: Heteroskedasticity consistent (White correction); *t*-statistics are in parentheses. A constant term is included in both regressions.
a. Cutoff point is in square brackets.
Source: Authors' calculations.

It has been argued that spillovers from joint ventures should be greater than those from FDI (establishment of majority-owned affiliates), since the foreign partner has less ability to control the behavior of the domestic partner, and the domestic partner has a greater incentive to pursue R&D itself (see, for example, Pack and Saggi 1997). In contrast, internalization through FDI should be better able to limit technology leakage. If this is indeed the case, then excluding joint ventures from the SPILLOVER measure of the share of foreign ownership and reestimating the equation should increase the magnitude of the negative spillovers. The evidence, however, does not support this argument (table 5). Instead, the magnitude of the spillover effect becomes smaller and statistically insignificant, although it remains negative. Thus excluding joint ventures has an offsetting effect. In part this reflects the fact that joint ventures have higher TFP growth than firms without foreign partnerships, which raises the average of the group without FDI. This finding illustrates that the initial result of negative spillovers may not be robust. Tests for spillovers with the methodology used here (and in the literature more generally) require some assurance that in distinguishing between two subsets of firms in an industry on the basis of whether or not there is majority foreign ownership (or more generally foreign links of some kind) one is not ignoring other important determinants of firm performance.

One such determinant likely to be important is firms' investment in improving their technology. The survey questionnaire reveals that joint ventures invested significantly more in training and new technologies than purely domestic firms. The technological ability and effort that many of the firms without foreign partners

Table 4. *Spillover Effects (Firms without Foreign Links)*

Dependent variable: Growth in sales	*Ordinary least squares estimation*	*Random-effects estimation*
Amemiya selection bias correction variable	Yes	Yes
Sector-specific returns to scale and markups	Yes	Yes
Spillovers (share of assets of firms with joint ventures and foreign direct investment)	–0.178**	–0.172**
	(3.125)	(2.054)
Dummy for 1994	0.002	0.002
	(0.215)	(0.178)
Dummy for 1995	–0.038**	–0.037**
	(–4.201)	(–3.934)
Dummy for 1996	–0.036**	–0.035**
	(–3.534)	(–3.642)
Observations	340	340
F-test	0.92	
Hausman test (random versus fixed effects)[a]		4.57 [14:45]
Adjusted R^2	0.887	0.843

* Significant at the 10 percent level.

** Significant at the 5 percent level.

Note: Heteroskedasticity consistent (White correction); *t*-statistics are in parentheses. A constant term is included in both regressions.

a. Cutoff point is in square brackets.

Source: Authors' calculations.

Table 5. *Testing for Spillover Effects (Firms without Foreign Direct Investment)*

Dependent variable: Growth in sales	*Ordinary least squares estimation*	*Random-effects estimation*
Amemiya selection bias correction variable	Yes	Yes
Sector-specific returns to scale and markups	Yes	Yes
Spillovers (share of assets of foreign affiliates in total assets of the sector)	–0.077	–0.074
	(1.425)	(1.218)
Dummy for 1994	0.003	0.002
	(0.897)	(0.178)
Dummy for 1995	–0.032**	–0.031**
	(–2.985)	(–2.257)
Dummy for 1996	–0.027*	–0.025
	(–1.847)	(–1.514)
Observations	431	431
F-test	0.91	
Hausman test (random versus fixed effects)[a]		4.13 [14.45]
Adjusted R^2	0.894	0.857

* Significant at the 10 percent level.

** Significant at the 5 percent level.

Note: Heteroskedasticity consistent (White correction); *t*-statistics are in parentheses. A constant term is included in both regressions.

a. Cutoff point is in square brackets.

Source: Authors' calculations.

expend may be too low to absorb spillovers when they occur, or the firms with foreign links may have absorbed a significant share of the available stock of labor with requisite skills. Also, given that FDI and joint ventures together account for a significant share of total assets, sales, and employment in the Czech Republic, the potential for positive spillovers may be significant among firms with foreign partnerships, such as from FDI firms to joint ventures and among joint ventures. This suggests that, if domestic firms were excluded from the sample, FDI would have a positive effect on firms with joint ventures. But the effect is not statistically significant (the *t*-statistic is 1.42, possibly reflecting the small sample size).

Finally, account also should be taken of the short time frame on which the study is focused. Spillovers may require more time to affect TFP growth rates. And, as mentioned, absorbing new techniques requires significant in-house technological effort, which may not be captured adequately by the production function methodology used. Clearly, further research is required.

V. Concluding Remarks

Firm-level data for the Czech Republic during 1992–96 suggest that foreign investment has the predicted positive impact on TFP growth of recipient firms. This result is robust to corrections for the sample selection bias that arises because foreign companies tend to invest in firms with above-average productivity. It is not surprising that foreign investment raises TFP growth (with a lag), given

that foreign investors transfer new technologies and knowledge to partner firms. FDI appears to have a greater impact on TFP growth than do joint ventures, suggesting that parent firms are transferring more knowledge (soft or hard) to affiliates than joint venture firms obtain from their partners.

Taken together, joint ventures and FDI appear to have a negative spillover effect on firms that do not have foreign partnerships. This effect is relatively large and statistically significant, and it cuts across industries. However, if we restrict attention to the impact of foreign-owned affiliates (FDI) on all other firms in an industry, the magnitude of the negative effect becomes much smaller and loses statistical significance. This result, in conjunction with the fact that joint ventures and FDI together account for a significant share of total output in many industries in the sample, suggests that further research is required to determine the extent to which knowledge diffuses from firms that have strong links to foreign firms to firms that do not have such relationships. Particularly important in this connection is exploring the extent of spillovers among joint ventures and between foreign affiliates and joint ventures. Insofar as joint ventures invest more in technological capacity (as is suggested by their training efforts), we would expect them to be better able to absorb and benefit from the diffusion of knowledge. The absence of such capacity may underlie the observed negative spillover effect. Longer time series and collection of data that measure firms' in-house technological efforts would help to identify the magnitude and determinants of technological spillovers.

Further analysis of the performance of Czech firms is necessary to see whether our results hold up for a larger sample of firms and in more recent years. Such data are being collected at the World Bank as part of a research project on knowledge transfer, and more robust results will emerge in the future.

REFERENCES

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Aitken, Brian, Gordon Hanson, and Ann Harrison. 1997. "Spillovers, Foreign Investment, and Export Behavior." *Journal of International Economics* 43(1–2):103–32.

Amemiya, Takeshi. 1984. "Tobit Models: A Survey." *Journal of Econometrics* 24(1):3–61.

Blomström, Magnus, and Ari Kokko. 1997. "How Foreign Investment Affects Host Countries." Policy Research Working Paper 1745. International Trade Department, World Bank, Washington, D.C. Processed.

Blomström, Magnus, Ari Kokko, and Mario Zejan. 1994. "Host Country Competition and Technology Transfer by Multinationals." *Weltwirtschaftliches Archiv* 130(3):521–33.

Blomström, Magnus, and Hakan Persson. 1983. "Foreign Investment and Spillover Efficiency in an Underdeveloped Economy: Evidence from the Mexican Manufacturing Industry." *World Development* 11(6):493–501.

Clerides, Sofronis, Saul Lach, and James Tybout. 1998. "Is 'Learning-by-Exporting' Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco." *Quarterly Journal of Economics* 114(3):452–75.

Coe, David, and Elhanan Helpman. 1995. "International R&D Spillovers." *European Economic Review* 39(5):214–42.

Coe, David, Elhanan Helpman, and Alexander Hoffmaister. 1997. "North-South R&D Spillovers." *Economic Journal* 107(440):134–49.

Czech Statistical Office. 1998. *Statistical Yearbook, 1997*. Prague.

Das, Saghamitra. 1987. "Externalities and Technology Transfer through Multinational Corporations: A Theoretical Analysis." *Journal of International Economics* 22(1–2):171–82.

Evenson, R. E., and Larry Westphal. 1995. "Technological Change and Technology Strategy." In Jere Behrman and T. N. Srinivasan, eds., *Handbook of Development Economics*. Vol. 3. Amsterdam: North-Holland.

Feenstra, Robert, James Markusen, and William Zeile. 1992. "Accounting for Growth with New Inputs: Theory and Evidence." *American Economic Review* 82(2):415–21.

Griliches, Zvi. 1992. "The Search for R&D Spillover." *Scandinavian Journal of Economics* 94(1):29–47.

Grossman, Gene, and Elhanan Helpman. 1995. "Technology and Trade." In Gene Grossman and Kenneth Rogoff, eds., *Handbook of International Economics*. Vol. 3. Amsterdam: North-Holland.

Haddad, Mona, and Ann Harrison. 1993. "Are There Positive Spillovers from Direct Foreign Investment? Evidence from Panel Data for Morocco." *Journal of Development Economics* 42(1):51–74.

Harrison, Ann. 1994. "Productivity, Imperfect Competition, and Trade Reform." *Journal of International Economics* 36(1):53–73.

———. 1996. "Determinants and Consequences of Foreign Investment in Three Developing Countries." In Mark Roberts and James Tybout, eds., *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*. Oxford: Oxford University Press.

Helleiner, Gerry. 1973. "Manufactured Exports from Less Developed Countries and Multinational Firms." *Economic Journal* 83(1):21–47.

Karshenas, Massoud, and Paul Stoneman. 1994. "Technological Diffusion." In P. Stoneman, ed., *Handbook of Economics of Innovation and Technological Change*. Oxford: Blackwell.

Keesing, Don, and Sanjaya Lall. 1992. "Marketing Manufactured Exports from Developing Countries: Learning Sequences and Public Support." In Gerry Helleiner, ed., *Trade Policy, Industrialization, and Development: New Perspectives*. Oxford: Clarendon Press.

Lall, Sanjaya. 1987. *Learning to Industrialize: The Acquisition of Technological Capabilities in India*. London: Macmillan.

———. 1992. "Technological Capabilities and Industrialization." *World Development* 20(2):165–86.

Markusen, James. 1995. "The Boundaries of the Multinational Enterprise and the Theory of International Trade." *Journal of Economic Perspectives* 9(1):169–89.

———. 1998. "Multilateral Rules on FDI: The Developing Countries Stake." International Trade Department, World Bank, Washington, D.C. Processed.

Nelson, Richard, and Edward Wolff. 1997. "Factors Behind Cross-Industry Differences in Technical Progress." *Structural Change and Economic Dynamics* 8(2):205–20.

Nelson, Richard, and Howard Pack. 1998. "The Asian Miracle and Modern Growth Theory." Policy Research Working Paper 1881. International Trade Department, World Bank, Washington, D.C. Processed.

Pack, Howard. 1987. *Productivity, Technology, and Industrial Development*. New York: Oxford University Press.

Pack, Howard, and Kamal Saggi. 1997. "Inflows of Foreign Technology and Indigenous Technological Development." *Review of Development Economics* 1(1):81–98.

Pack, Howard, and Larry Westphal. 1986. "Industrial Strategy and Technological Change." *Journal of Development Economics* 22(1):87–128.

Ray, G. F. 1964. *The Diffusion of Mature Technologies*. Cambridge: Cambridge University Press.

Smarzynska, Beata. 1998. "Composition of FDI and Protection of Intellectual Property Rights in Transition Economies." Economics Department, Yale University, New Haven, Conn. Processed.

Wang, Jian-Ye, and Magnus Blomström. 1992. "Foreign Investment and Technology Transfer: A Simple Model." *European Economic Review* 36(1):137–55.

Westphal, Larry, Yung Rhee, and Garry Pursell. 1981. *Korean Industrial Competence: Where It Came From*. World Bank Staff Working Paper 469. Washington, D.C.: World Bank.

THE WORLD BANK ECONOMIC REVIEW, VOL. 14, NO. 1: 65-90

Productivity and Turnover in the Export Market: Micro-level Evidence from the Republic of Korea and Taiwan (China)

Bee Yan Aw, Sukkyun Chung, and Mark J. Roberts

Widespread empirical evidence indicates that exporting producers have higher productivity than nonexporters, although the reasons why are unclear. Some analysts argue that exporters acquire knowledge of new production methods, inputs, and product designs from their international contacts, and with this knowledge they achieve higher productivity than their more insulated domestic counterparts. Others argue that the higher productivity of exporters reflects the self-selection of more efficient producers into a highly competitive export market. This article analyzes the link between a producer's total factor productivity and its decision to participate in the export market, using manufacturing data from the Republic of Korea and Taiwan (China).

Differences are found between these two economies in the importance of selection and learning. In Taiwan (China) transitions of plants into and out of the export market reflect systematic variations in productivity as predicted by self-selection models. In Korea there are no significant changes in productivity following entry or exit from the export market that are consistent with learning from exporting. A comparison of the two economies suggests that in Korea factors other than production efficiency are more prominent determinants of the export decision.

Over the past three decades the Republic of Korea and Taiwan (China) have achieved high and sustained rates of growth in output and income. Although high savings rates and substantial investments in new capital equipment have contributed to their success, it is impossible to ignore the role of the export market. At a minimum, the ability to export has allowed manufacturers to specialize in a range of products and to increase their output levels far beyond what the size of their domestic market could support. Some economists attribute the success of these economies to the role of exports in serving as a conduit for technology transfer from abroad and in generating technological spillovers to the rest of the economy. Case studies and empirical evidence show that exporting firms or plants

Bee Yan Aw is with the Department of Economics at Pennsylvania State University, Sukkyun Chung is with the Ministry of Finance and Economy of the Republic of Korea, and Mark J. Roberts is with the Department of Economics at Pennsylvania State University and the National Bureau of Economic Research (NBER). Their e-mail addresses are *byr@psu.edu*, *schung@mic.go.kr*, and *mroberts@psu.edu*. The authors appreciate the helpful comments of J. David Richardson, Jim Tybout, and participants at the Trade and Technology Diffusion conference at the Fondazione Eni Enrico Mattei, Milan, Italy, and the NBER Summer Productivity Institute.

are more efficient than their counterparts that sell primarily in the domestic market. This belief that export activity generates cumulative productivity benefits is often cited as an argument for the active promotion of exports in many developing countries.

The empirical finding that exporters are more productive than nonexporters is widespread and robust, but only two recent papers have addressed the more complex issue of whether exports play a causal role in generating higher productivity. Clerides, Lach, and Tybout (1998) examine this issue using manufacturing data for Colombia, Mexico, and Morocco. Bernard and Jensen (1999) focus on manufacturers in the United States. Both articles examine two alternative explanations linking productivity and exporting. The first explanation holds that exporters learn from their contacts in the export market, and as a result they adopt better production methods and achieve higher productivity. The second says that the higher productivity of exporting firms may reflect the self-selection of more efficient producers into a highly competitive export market. Both studies find that the self-selection of more efficient producers into the export market is an important part of the story and that there is little evidence of efficiency gains that could reflect the learning that accrues from exporting.

In this article we study the link between a producer's total factor productivity (TFP) and its choice to participate in the export market, using micro-data collected in manufacturing censuses in Korea and Taiwan (China). We specifically focus on the relationship between productivity and the movement of producers into and out of the export market. Productivity differences between producers with different transition patterns, rather than just different exporting status, are crucial to separating the selection and learning explanations. If differences in productivity predate a producer's movement into or out of the export market so that nonexporters with high productivity tend to enter whereas low-productivity exporters tend to exit, then self-selection forces are at work. In contrast, if differences in productivity follow a producer's transition into or out of the export market, then learning-by-doing forces are at work. Producers that enter the export market should subsequently have greater productivity changes than producers who do not enter. Likewise, producers that exit the export market should begin to lag behind their counterparts that remain in.

I. The Relationship between Productivity and Exporting

A large body of empirical evidence demonstrates that firms that participate in the export market perform better than other firms in terms of productivity, size, length of survival, and wages paid.[1] The literature proposes that at least two

1. Several papers examine the export-productivity relationship at the micro-level (see Aw and Hwang 1995; Aw and Batra 1998; Chen and Tang 1987; Haddad 1993; Handoussa, Nishimizu, and Page 1986; and Tybout and Westbrook 1995). Aw and Batra (1999) and Bernard and Jensen (1995) examine the relationships among exports, firm size, and wages. Richardson and Rindal (1995, 1996) summarize the empirical evidence for a wide range of firm characteristics that are correlated with the exporting activity.

mechanisms explain the positive correlation between exporting and productivity. First, the correlation may simply reflect the fact that only the most productive firms survive in the highly competitive export market. If the fixed costs of selling are higher in the export market than in the domestic market or if output prices are lower, only firms with high productivity will find it profitable to enter the export market in the first place. Exporters whose productivity declines will be forced to exit. We refer to this as the self-selection hypothesis.

Second, the correlation may reflect productivity improvements that result from knowledge and expertise that the firm gains as a direct result of its experience in the export market. Some analysts argue that firms that participate in the export market gain access to technical expertise from their buyers, including both new product designs and production methods, that nonexporters do not have. This phenomenon of learning by exporting may be particularly relevant for countries in East Asia.[2]

Both mechanisms are plausible, but their actual importance most likely varies across countries and industries. Different rates of product and process innovation alter the possibilities for learning and the nature of trade policy, which can affect the strength of market selection forces. Clerides, Lach, and Tybout (1998) and Bernard and Jensen (1999) find clear evidence of the importance of self-selection among exporters in the 1980s. Firms that become exporters perform more efficiently prior to entry than their nonexporting counterparts. In addition, both studies find little evidence of efficiency gains that could reflect learning by exporting. Clerides, Lach, and Tybout find that a producer's exporting history does not significantly alter current production costs. Bernard and Jensen find that future productivity growth is not significantly higher for plants that currently export. They do find that employment growth and the probability of survival are higher for exporting plants. This pattern can reflect the same underlying differences in efficiency that lead to self-selection into the export market and does not necessarily reflect improvements that follow as a result of exporting experience. Overall, the evidence weighs heavily on self-selection of the more efficient firms into the export market as the main source of the productivity differences between exporters and nonexporters.

Theoretical Framework

Several recent theoretical models of industry dynamics explain the divergent paths of growth and failure that characterize micro-data on individual producers. All of these models begin with the assumption that producers within the same industry differ in their productive efficiency and are subject to idiosyncratic shocks or uncertainty. Differences in the evolution of their productivity over time, in turn, lead producers to make different decisions regarding entry, growth, and exit. The source of uncertainty differs across models. Jovanovic (1982) empha-

2. See Evenson and Westphal (1995); Grossman and Helpman (1991); Rhee, Ross-Larson, and Pursell (1984); and World Bank (1993) for discussions and evidence on the role of buyers in providing technical expertise.

sizes firms' uncertainty about their own productivity levels, Lambson (1991) focuses on uncertain future market conditions, Hopenhayn (1992) emphasizes randomness in productivity changes over time, and Ericson and Pakes (1995) model uncertainty in the return to firm investments.

To organize our empirical analysis of productivity and the decision to export, we rely on Hopenhayn's (1992) model. Although not specific to the export market, Hopenhayn's model shows how firms with different levels of productivity make different decisions to enter, exit, or remain in a market. It allows us to identify how self-selection will lead to differences in productivity among these three cohorts of firms.

Hopenhayn models a market that is composed of a large number of price-taking firms that produce a homogeneous output. Firms differ in their efficiency, with each firm's output depending on a random productivity shock, θ, which follows a Markov process that is independent across firms. The distribution of future productivity is described by the distribution function $F(\theta_{t+1} \mid \theta_t)$, which is assumed to be strictly decreasing in θ_t. This assumption implies that, relative to a firm with low θ_t, a firm with high productivity in year t has a larger probability of having high productivity in year $t + 1$.

In each period, before the new productivity shock is observed, incumbent firms may choose to exit the market or to remain in the market and pay a fixed cost. Following that decision, they observe their productivity shock and choose their output level for that period. Potential entrants may choose to enter by paying a sunk entry cost, after which they draw their initial productivity level from a common distribution function, $G(\theta)$, and choose their output level. Output prices are determined competitively to equate demand and supply. The model produces two key endogenous variables: the flow of entrants into the market each period and the minimum productivity level needed for an incumbent firm to remain in the market. This productivity level, which we denote X_t, is the lowest level of productivity that will generate positive expected profits for the firm over future periods.

This model makes predictions about differences in the average productivity of entering, exiting, and surviving producers. Hopenhayn demonstrates that firms will exit the market after period t if $\theta_t < X_t$. The current-period productivity, θ_t, which the firm observes, determines the likely future trajectory of productivity through the distribution function $F(\theta_{t+1} \mid \theta_t)$. Firms with $\theta_t < X_t$ expect low future profit streams and exit after period t, while firms with $\theta_t > X_t$ remain in the market. The model implies that exiting firms are concentrated among the least-productive firms. We can test this implication empirically for the export market by examining exporters in period t and looking for systematic differences in productivity between the group that continues exporting in $t + 1$ and the group that exits.

The model allows us to compare the productivity of a cohort of new entrants with the productivity of cohorts of older surviving producers. The distribution function for initial productivity, $G(\theta)$, determines the productivity of the new

entrants. The productivity of older cohorts reflects the failure of the least-productive members over time and the random changes in the survivors' productivity over time. Hopenhayn demonstrates that if $F(\theta_{t+1} \mid \theta_t)$ is strictly decreasing in θ_t, then the productivity distribution of any surviving cohort stochastically dominates the productivity distribution of the entering cohort. We can examine this empirically by comparing the productivity of new exporters and incumbent exporters at a point in time.

In the formal model all firms base their entry decision only on knowledge of the distribution of initial productivity $G(\theta)$ and its evolution over time $F(\theta_{t+1} \mid \theta_t)$, and not on information about their own productivity level. Many potential entrants to the export market currently produce in the domestic market and thus have better information on their likely productivity after entry than firms with no prior experience. Therefore, domestic producers with high productivity in year t would be more likely to enter the export market than low-productivity domestic producers. We can examine this empirically by focusing on firms that initially produce only in the domestic market. We ask whether the firms that subsequently enter the export market have higher initial productivity than those that remain specialized in the domestic market.

Although Hopenhayn's model clarifies the important role of firm heterogeneity and self-selection in generating flows of firms into and out of a market, it does not incorporate the idea that productivity may change following entry, as described in the learning-by-exporting hypothesis. Clerides, Lach, and Tybout (1998) include this possibility in a model of a domestic firm's decision to diversify into the export market. The firm makes the decision to enter by comparing expected future profits, which depend on the firm's current and future productivity, with the sunk costs of entry. To incorporate the effect of learning by exporting, Clerides, Lach, and Tybout make a firm's current productivity depend on prior export experience.

Clerides, Lach, and Tybout include a set of simulation results that provide useful insights into how selection and learning interact. They find that firms that enter or remain in the export market always have higher productivity than firms that stop exporting or that remain only in the domestic market. Firms that enter the export market also have higher productivity prior to entry than firms that produce only in the domestic market. Both of these patterns arise because firms self-select into the export market based on their current productivity. Adding learning by exporting to the framework widens the gap between the productivity of firms that enter the export market and those that do not. We can examine this pattern empirically by comparing the pre- and post-entry productivity differentials between entrants and nonentrants.

Empirical Implications in Micro-Data from Korea and Taiwan (China)

We analyze a data set that includes information collected as part of the manufacturing censuses in Korea and Taiwan (China). Appendix A gives a description of the data. In the case of Taiwan (China) observations are at the firm level for

the census years 1981, 1986, and 1991. However, for the industries we study, between 80 and 90 percent of all Taiwanese firms are single-plant producers; therefore, the distinction between plant and firm is not as important as it is in many industrial countries. In the case of Korea we have plant-level observations for the years 1983, 1988, and 1993. For simplicity we refer to the data as plant-level for both economies, even though only firm-level information is available for Taiwan (China).

The plant-level observations have been matched over time so that it is possible to identify entering and exiting producers in each census year. In addition, plant-level exports are reported for all three years in Korea and for 1986 and 1991 in Taiwan (China).[3] The data allow us to classify each producer as a nonexporter, an entrant to the export market, an incumbent exporter, or a plant that has exited the export market between each pair of years.

The data set contains information on output and inputs of capital, labor, and raw materials that allows us to construct an index of TFP for each plant. Appendix B gives details on the construction of the productivity index. Although we have time observations for only three years, the fact that the censuses are taken at five-year intervals provides an advantage over a data set with a small number of observations for consecutive years. The longer five-year time period reduces the importance of transitory shocks, cyclical fluctuations, and measurement errors in our productivity comparisons. It also makes it more likely that we will observe long-term changes in productivity than would comparisons based on data of higher frequency.

Given our focus on the role of the export market as a source of knowledge and productivity differentials, we restrict our attention to the five two-digit industries that have a major export role in both Taiwan (China) and Korea—textiles, apparel, plastics, electrical machinery and electronics, and transportation equipment. These industries have the highest export participation rates in the manufacturing sector. (The participation rate is the share of plants that export.) In Taiwan (China) the export participation rate ranges from 26 percent in transportation equipment to 41 percent in electrical machinery and electronics. In Korea the participation rate ranges from 13 percent in apparel to 26 percent in electronics. In both economies the five industries account for more than half of total manufacturing exports in 1986 and 1988, respectively.

To separate the effects of selection and learning, Clerides, Lach, and Tybout (1998) use plant-level panel data with a relatively long time-series component to estimate a two-equation model consisting of the plant's decision to participate in the export market and the plant's cost function. The micro-data for Korea and Taiwan (China) do not have sufficient time-series observations to allow us to use this approach. Our basic empirical strategy, which is similar to that of

3. For Taiwan (China) we observe the level of exports and domestic sales for each firm. For Korea we have the value of plant sales and a set of categorical variables indicating whether the plant's export-sales ratio is high (more than 0.75), moderate (0.25–0.75), low (positive but less than 0.25), or zero.

Bernard and Jensen (1999), is to compare the average productivity of groups of plants that have undergone different transition patterns. As indicated by Hopenhayn's model, self-selection implies that a plant's productivity level in period t should be a determinant of export market participation in year $t + 1$. The learning-by-exporting explanation implies that initial productivity differences between plants that select into the market and those that do not should widen following entry or as firms accumulate more experience in the export market. To isolate this effect, we focus on changes in the differential between periods t and $t + 1$ for exporters and nonexporters.

Aside from learning by exporting, there are several explanations as to why the productivity of exporters changes more than the productivity of nonexporters over time. If entry into the export market allows plants to expand their output and take advantage of economies of scale in production, then exporters will have larger increases in productivity than nonexporters. In general, any factor that results in positive serial correlation in the shocks to plant-level productivity will generate this outcome. Plants experiencing positive productivity shocks are more likely to find it profitable to enter the export market. If these positive shocks continue over time, the productivity of exporters and nonexporters will continue to diverge. With our data we cannot distinguish these alternative explanations. However, the finding that productivity differences between exporters and nonexporters do not diverge following entry would be inconsistent with any of these explanations, including learning by exporting.

To clarify the comparisons, we define four groups of plants based on their participation in the export market in two adjoining years of data (table 1). We make five different comparisons based on these four groups.

In section II we compare the productivity of exporters and nonexporters in each year in order to confirm the positive cross-sectional correlation between exporting and productivity. Then, in section III, we compare the productivity of the four transition groups in the same year in order to see if the decision to participate in the export market reflects plant productivity. In both the first and second comparisons we use all producers operating in the year of interest. Thus we include in comparisons for year t failing plants that exit production entirely after year t, and we include in comparisons for year $t + 1$ new plants that enter production after year t.

In the last three comparisons we use the subset of plants that operate in both years so that we can compare improvements or declines in productivity with

Table 1. *Definitions of Four Groups of Plants Based on Their Participation in the Export Market*

Group	*Plant status*	*Year t*	*Year t + 1*
1	Stay out	Does not export	Does not export
2	Enter	Does not export	Exports
3	Exit	Exports	Does not export
4	Stay in	Exports	Exports

experience in the export market. In section IV we look at nonexporters in year t (groups 1 and 2 in table 1) and compare productivity in years t and $t + 1$. If market selection is important, the productivity of the entrants (group 2) should exceed that of the plants that stay out of the export market (group 1) in year t. Comparing the productivities of the two groups in year $t + 1$ reveals whether the initial differential narrows, widens, or remains unchanged after group 2 has gained some experience in the export market. If learning is important, this differential should widen.

To determine whether or not productivity differentials persist following exit, the fourth comparison, in section V, looks at productivity in years t and $t + 1$ for groups 3 and 4. Plants in these groups begin in the export market and follow different paths over time. If market selection is important, then exit from the export market should be concentrated in plants with lower productivity. If exporting brings additional benefits, then the productivity differential should widen in period $t + 1$ between the group that stays in the export market and the group that exits.

With the fifth comparison, in section VI, we look at whether or not exporters follow different productivity paths than nonexporters over time. We compare the productivity of groups 1 and 4 in years t and $t + 1$. If the export market facilitates the accumulation of knowledge over time, productivity levels between the two groups should increasingly diverge. As a further refinement to this comparison, we look at whether improvements over time accrue to new producers, the group most likely to benefit from learning effects. To do this, we compare productivity in years t and $t + 1$ for the producers in groups 1 and 4 that first begin operating in year t.

II. Productivity Differentials between Exporting and Nonexporting Plants

We begin by summarizing the cross-sectional differences in average productivity between plants that sell in the export market and those that operate solely in the domestic market (table 2). For example, in the textile industry in Taiwan (China) in 1986, exporting plants had 27.6 percent higher TFP levels than nonexporting plants. Across the five industries in Taiwan (China) average TFP levels were higher for exporters than for nonexporters by between 11.8 percent (electrical machinery in 1986) and 27.6 percent (textiles in 1986). All of the differences in means are statistically significant. The data for Korea similarly show higher productivity among exporting plants. The average productivity difference between exporters and nonexporters varies between 3.9 percent (electrical machinery in 1988) and 31.1 percent (textiles in 1983), and all the differences are statistically significant.[4]

4. In both economies these differences are smaller in the newer, higher-technology industries of electronics and transportation. Pack (1992) argues that worker mobility is one way that knowledge gained in the export market can diffuse to other producers. If labor market turnover is higher in industries that use rapidly changing technologies, then the positive spillovers from exporting to nonexporting plants may be higher in these industries. This transmission of knowledge through worker movement would result in smaller average productivity differentials between exporters and nonexporters.

Table 2. *Productivity Differences between Exporters and Nonexporters and the Number of Exporting and Nonexporting Firms in the Republic of Korea and Taiwan (China), 1980s and 1990s*

	Korea			*Taiwan (China)*	
Industry and indicator	*1983*	*1988*	*1993*	*1986*	*1991*
Textiles					
Percentage difference in average productivity[a]	0.311	0.234	0.231	0.276	0.186
	(0.017)	(0.014)	(0.014)	(0.010)	(0.010)
Exporters	510	874	1,163	1,231	946
Nonexporters	1,368	1,767	2,352	2,039	2,589
Apparel					
Percentage difference in average productivity[a]	0.189	0.153	0.199	0.247	0.196
	(0.022)	(0.018)	(0.019)	(0.011)	(0.013)
Exporters	257	499	479	809	571
Nonexporters	1,479	1,852	2,212	1,171	1,465
Plastics					
Percentage difference in average productivity[a]	0.148	0.097	0.071	0.166	0.151
	(0.027)	(0.016)	(0.014)	(0.006)	(0.007)
Exporters	193	481	572	1,806	1,497
Nonexporters	1,171	2,109	3,563	4,811	7,470
Electrical machinery and electronics					
Percentage difference in average productivity[a]	0.068	0.039	0.045	0.118	0.145
	(0.021)	(0.013)	(0.011)	(0.007)	(0.006)
Exporters	385	880	1,149	2,024	2,347
Nonexporters	933	1,917	3,735	3,354	5,703
Transportation equipment					
Percentage difference in average productivity[a]	0.140	0.110	0.094	0.126	0.153
	(0.036)	(0.021)	(0.017)	(0.010)	(0.011)
Exporters	98	248	266	606	678
Nonexporters	507	1,003	2,045	1,751	2,565

a. The values show the percentage by which TFP is higher in exporting than in nonexporting firms. The standard errors of the differences are in parentheses.

Source: Authors' calculations.

The simple comparison of average productivity in table 2 clearly indicates the higher productivity of exporters relative to nonexporters in both countries.[5] The results in table 2 mirror the findings for virtually every other country for which micro-level productivity comparisons have been done. But the underlying causal mechanism is unclear. If the domestic market is limited in size, then firms can benefit from entering the larger export market. However, higher levels of com-

5. See Tybout (1996) for a summary of the empirical literature on productivity differences among firms.

petition in the world market or higher fixed costs associated with selling in the export market mean lower profit streams per unit, so that only the more efficient firms will enter and survive in the export market. Alternatively, if firms already in the export market can take advantage of scale economies or acquire knowledge of new technologies that foster learning, this will be reflected in higher productivity for exporters.

If these externalities from exporting exist, it is very likely that they are higher the greater is the degree of exposure to the export market. Therefore, we look at whether the productivity differential is an increasing function of the share of plant output that is exported and whether the differential is independent of the degree of exposure. Table 3 reports the results of regressions of plant productivity on year and export intensity dummies for each country and industry. The intercept represents the plants that do not export. The remaining coefficients measure the percentage difference in productivity between nonexporters and plants with low export intensity (less than 25 percent of production exported), moderate intensity (25 to 75 percent), and high intensity (more than 75 percent). The positive and significant coefficients on the export intensity dummies for both countries clearly indicate higher levels of productivity for exporting firms relative to nonexporters, as demonstrated in table 2.

The new finding is that differences in average productivity across groups of plants with different export intensities are very small, particularly when compared with the differences between exporters and nonexporters (table 2). The data for Taiwan (China) show that it is not possible to reject the hypothesis that average productivity is the same across all three categories of export intensity for the textile and electrical machinery industries. For the other three industries we cannot reject the hypothesis that two of the three groups have equal average productivity. In addition, there is no consistent movement in the level of average productivity across intensity categories. For two industries productivity falls moving from low to high export intensity; for three industries it increases. Except for the apparel industry, the direction of change within industries is not monotonic across intensity categories.

The data for Korea show similar patterns. For three of the five industries we do not reject the hypothesis that the three export categories have the same average productivity. In the textile and transportation industries the evidence indicates that the plants that export at least 75 percent of their output do have higher productivity. Average productivity among Korean textile plants that export less than one-quarter of their output is 18.8 percent higher than that of nonexporters, and this differential rises to 28.1 percent for plants that export at least three-quarters of their output. In transportation equipment, exporters in the low-intensity category are 9.4 percent more productive than nonexporters, and the differential rises to 20.2 percent for the high-intensity category.

Overall, the cross-sectional results in tables 2 and 3 indicate that being an exporter, per se, signals higher productivity in every case, but the percentage of the plant's output that is exported has little systematic effect on productivity for

Table 3. *The Impact of Export Intensity on Plant Productivity*

Country and industry	*Intercept*	*Export intensity*[a] *Low*	*Medium*	*High*	*Test results*[b]
Korea					
Textiles	−0.118*	0.188*	0.228*	0.281*	2
	(0.009)	(0.018)	(0.014)	(0.011)	
Apparel	−0.068*	0.242*	0.176*	0.173*	1, 2, 3
	(0.009)	(0.037)	(0.030)	(0.013)	
Plastics	−0.067*	0.092*	0.084*	0.111*	1, 2, 3
	(0.009)	(0.014)	(0.016)	(0.020)	
Electrical machinery and electronics	−0.079*	0.058*	0.024*	0.055*	1, 2, 3
	(0.009)	(0.013)	(0.013)	(0.012)	
Transportation equipment	−0.070*	0.094*	0.085*	0.202*	2
	(0.013)	(0.017)	(0.024)	(0.033)	
Taiwan (China)					
Textiles	−0.012*	0.236*	0.212*	0.244*	1, 2
	(0.005)	(0.014)	(0.012)	(0.009)	
Apparel	−0.142*	0.181*	0.193*	0.233*	2
	(0.007)	(0.027)	(0.018)	(0.009)	
Plastics	0.012*	0.145*	0.141*	0.170*	2
	(0.003)	(0.010)	(0.009)	(0.006)	
Electrical machinery and electronics	−0.007	0.145*	0.129*	0.131*	1, 2, 3
	(0.004)	(0.009)	(0.007)	(0.006)	
Transportation equipment	−0.140*	0.179*	0.121*	0.133*	3
	(0.005)	(0.015)	(0.014)	(0.010)	

* Significant at the 5 percent level.

Note: All regressions include year dummy variables. Standard errors are in parentheses.

a. Export intensity is low if the export share is greater than 0 and less than or equal to 0.25, medium if the export share is greater than 0.25 and less than or equal to 0.75, and high if the export share is greater than 0.75.

b. Test results are coded as follows (all are for the 5 percent level of significance): 1, do not reject the equality of all three export intensity coefficients; 2, do not reject the equality of the low and medium export intensity coefficients; and 3, do not reject the equality of the medium and high export intensity coefficients.

Source: Authors' calculations.

most of the industries. That export intensity may not be a good measure of the extent of knowledge that a plant gains from foreign sources could explain this result. An exporter has access to a pool of new ideas that is more likely to be a function of the exporter's number of foreign purchasers or contacts, rather than the percentage of output that it exports. Unfortunately, we do not have any information on the buyers of each plant's output; we can distinguish only whether the plant has some foreign contact or no foreign contact based on its total volume of exports.

III. Productivity Differentials between Transition Groups

To analyze productivity differentials between transition groups, we exploit the time-series aspects of the data and combine information on the transition

patterns of plants in the export market with the cross-sectional productivity distribution. Our regression results compare the productivity of all plants in year $t + 1$ based on whether they enter or exit the export market (table 4). The plants that do not export in either year (group 1 in table 1) make up the base category.

For Taiwan (China) there is an identical ranking of categories for all five exporting industries. The group with the lowest average productivity stays out of the export market in both years. Exiting plants have average productivity levels that are 4.4 to 10.3 percent higher than plants that have never exported. Entrants are 13.3 to 18.9 percent more productive than nonexporters. Finally, plants that remain in the export market are 16.7 to 22.3 percent more productive than nonexporters.

Table 4. *The Impact of Transitions into or out of the Export Market on Plant Productivity*

		Differential for plants that[a]		
Country and industry	*Intercept*	*Exit the export market*	*Enter the export market*	*Remain in the export market*
Korea				
Textiles	–0.112*	0.115*	0.240*	0.209*
	(0.008)	(0.032)	(0.012)	(0.017)
Apparel	–0.061*	0.131*	0.186*	0.121*
	(0.008)	(0.048)	(0.015)	(0.030)
Plastics	–0.040*	–0.004	0.077*	0.102*
	(0.006)	(0.028)	(0.012)	(0.022)
Electrical machinery and electronics	–0.025*	–0.032	0.037*	0.056*
	(0.007)	(0.026)	(0.009)	(0.016)
Transportation equipment	–0.022*	–0.018	0.086*	0.149*
	(0.009)	(0.038)	(0.016)	(0.029)
Taiwan (China)				
Textiles	0.150*	0.103*	0.173*	0.223*
	(0.005)	(0.021)	(0.012)	(0.014)
Apparel	–0.018*	0.064*	0.189*	0.219*
	(0.007)	(0.028)	(0.015)	(0.020)
Plastics	0.069*	0.082*	0.138*	0.196*
	(0.003)	(0.014)	(0.008)	(0.012)
Electrical machinery and electronics	0.186*	0.044*	0.138*	0.167*
	(0.003)	(0.014)	(0.007)	(0.009)
Transportation equipment	–0.205*	0.080*	0.133*	0.211*
	(0.005)	(0.023)	(0.013)	(0.018)

*Significant at the 5 percent level.

Note: All regressions include year dummy variables. Standard errors are in parentheses.

a. The percentage difference in average productivity between the given category and firms that do not export (group 1).

Source: Authors' calculations.

Exiting firms are between 11.4 and 15.5 percent less productive than plants that remain in the export market (columns 2 and 4 of table 4). In addition, entrants are between 2.9 and 7.8 percent less productive than experienced exporters (columns 3 and 4 of table 4). Both patterns are consistent with the model of self-selection outlined in section I.

The patterns for Korea differ in some systematic ways from the results for Taiwan (China). First, in three industries—plastics, electrical machinery, and transportation equipment—the average productivity of plants that exit the export market is not significantly different from that of plants with no export market experience. Second, in two cases—textiles and apparel—entrants are more productive than incumbent exporters. Third, in the apparel industry, exiting and surviving plants have nearly the same average productivity (0.131 and 0.121 percent, respectively). All of these patterns indicate that, relative to Taiwan (China), differences in productivity in Korea are not as closely related to transitions into or out of the export market.

An additional refinement we make is to further divide the plants in year $t + 1$ into new plants, those that first appear in production in year $t + 1$, and old plants, those that were present in year t. In Taiwan (China) the differences between the two groups are minimal, with one exception. In the apparel industry, among new plants entering the export market, average productivity is 10 percent higher than that of old plants entering the export market.

In Korea two industries have substantial differences. New plants in textiles are approximately 16 percent more productive than old plants, and this differential holds for both exporters and nonexporters. In apparel new plants that enter the export market are on average 14.3 percent more productive than old plants that begin exporting, but there is no difference between new plants and plants that do not export. The productivity difference between new and old plants can reflect the adoption of different technologies in the new plants. Because this differential is observed for both exporters and nonexporters in Korean textiles, it is unlikely that exporting is the conduit for the improvement in technology. However, for the apparel industries, only new exporting plants have higher productivity. This result raises the possibility that knowledge transfers resulting from contacts with foreign buyers could be the mechanism at work.

Our finding that Taiwanese plants that exit the export market have higher average productivity than nonexporters differs from the findings of Clerides, Lach, and Tybout (1998) and Bernard and Jensen (1999). Both studies find that plants exiting the export market are among the worst performers. One explanation may be that the sunk costs involved in reentering the export market in Taiwan (China) are sufficiently low that plants do not hesitate to exit in the face of low productivity. By contrast, if other countries have higher entry costs, producers may prefer to remain in the export market in the face of low productivity and profits in order to wait and see if productivity improves. When sunk costs are high, firms face a high option value of remaining in operation in order to avoid reentry costs.

Only plants with very low productivity will choose to exit when entry costs are high.[6]

IV. Productivity Differentials between Entrants and Nonentrants

We now focus on entrants and nonentrants to the export market (groups 1 and 2 in table 1), comparing their average productivity in t and $t + 1$ (table 5). In every industry in Taiwan (China) plants that choose to enter the export market have significantly higher average productivity, prior to entry, than plants that choose to stay out. The differential varies from 4.8 percent in electrical machinery to 14.8 percent in apparel. This result is consistent with the self-selection hypothesis. The initial differential between the two groups of plants widens after entry in three of the industries—textiles, plastics, and electrical machinery (column 3 of table 5). The increase in the productivity differential is between 6.0 and 8.3 percent. In the other two industries the change in the differential following entry is not statistically significant. For textiles, plastics, and electrical machinery the widening productivity differential is consistent with the learning-by-exporting hypothesis. However, as noted, this hypothesis cannot be distinguished from other explanations that would generate positive serial correlation in productivity.

Overall, the results for Taiwan (China) indicate that, among continuing plants, productivity differences prior to entry are correlated with the entry decision. In several industries the plants that choose to enter continue to increase their productivity relative to nonentrants in the years following entry. The importance of self-selection into the export market is similar to the findings of Bernard and Jensen (1999) and Clerides, Lach, and Tybout (1998). However, none of the countries they examine shows evidence similar to what we find on the feedback effect on productivity from participating in the export market.

The data for Korea show a different pattern. The positive coefficients for all industries indicate that plants that choose to enter have higher productivity prior to entry than do nonexporters (column 2 of table 5). However, the difference is not statistically significant in two of the five industries. The productivity differential is statistically significant in textiles, where the differential is 17.6 percent; in transportation equipment, where it is 11.5 percent; and in plastics, where the differential is 5.8 percent. In addition, the differential between entrants and nonentrants widens following entry, but the change is not statistically significant in all five industries (column 3 of table 5). Thus the statistical evidence in support of both the self-selection and learning hypoth-

6. Roberts and Tybout (1997) develop the empirical implications of sunk entry costs on plant-level export participation. They find that sunk entry costs are an important determinant of exporting among Colombian manufacturing plants. They cite the absence of a well-developed export trading sector as one source of high entry costs. Campa (1998) finds that sunk exporting costs are also important for firms in the Spanish manufacturing sector. Levy (1991) argues that the well-developed network of trading firms in Taiwan (China) acts to lower the entry costs of new exporters.

Table 5. *Average Productivity Differences between Entrants and Nonentrants to the Export Market*

Country and industry	*Intercept*	*Entering firm differential, pre-entry*	*Change in differential, post-entry*
Korea			
Textiles	–0.143*	0.176*	0.059
	(0.013)	(0.026)	(0.033)
Apparel	0.014	0.036	0.111
	(0.019)	(0.052)	(0.074)
Plastics	–0.024	0.058*	0.008
	(0.014)	(0.027)	(0.038)
Electrical machinery and electronics	–0.006	0.016	0.027
	(0.019)	(0.026)	(0.036)
Transportation equipment	–0.036	0.115*	0.002
	(0.024)	(0.039)	(0.053)
Taiwan (China)			
Textiles	–0.007	0.060*	0.083*
	(0.010)	(0.026)	(0.037)
Apparel	–0.163*	0.148*	–0.026
	(0.013)	(0.044)	(0.062)
Plastics	0.018*	0.076*	0.061*
	(0.005)	(0.015)	(0.021)
Electrical machinery and electronics	0.007	0.048*	0.060*
	(0.008)	(0.016)	(0.023)
Transportation equipment	–0.134*	0.099*	0.025
	(0.010)	(0.028)	(0.039)

*Significant at the 5 percent level.
Note: All regressions include year dummy variables. Standard errors are in parentheses.
Source: Authors' calculations.

eses is much weaker for Korea than for Taiwan (China). The signs of the estimated coefficients are consistent with both effects, but the results are not generally statistically significant.

V. Productivity Differentials between Exiting and Surviving firms

For Taiwan (China) there is a difference in productivity between plants that exit the export market and those that remain (groups 3 and 4 in table 1). Plants that exit the export market after year t are less productive (in year t) than their counterparts that continue exporting, as indicated by the negative and significant coefficients (column 2 of table 6). The productivity gap varies from 6.2 to 13.1 percent. This result is consistent with the self-selection hypothesis.

Further, in four of the five industries in Taiwan (China), the plants that exit the export market fall further behind the exporting plants in the years following exit, as indicated by the negative regression coefficients (column 3 of table 6). This widening of the productivity differential between exporting and nonexporting

Table 6. *Average Productivity Differences between Plants that Exit the Export Market and Continuing Exporters*

Country and industry	*Intercept*	*Exiting firm differential, pre-exit*	*Change in differential, post-exit*
Korea			
Textiles	0.200*	–0.083*	–0.001
	(0.017)	(0.023)	(0.034)
Apparel	0.125*	0.076	–0.076
	(0.032)	(0.042)	(0.058)
Plastics	0.230*	–0.041	–0.047
	(0.032)	(0.032)	(0.044)
Electrical machinery and electronics	0.068*	–0.090*	0.012
	(0.023)	(0.027)	(0.037)
Transportation equipment	0.153*	–0.053	–0.091
	(0.035)	(0.039)	(0.053)
Taiwan (China)			
Textiles	0.302*	–0.121*	0.001
	(0.012)	(0.022)	(0.031)
Apparel	0.144*	–0.131*	–0.024
	(0.016)	(0.029)	(0.040)
Plastics	0.209*	–0.070*	–0.045*
	(0.010)	(0.016)	(0.022)
Electrical machinery and electronics	0.152*	–0.069*	–0.054*
	(0.007)	(0.014)	(0.019)
Transportation equipment	0.030*	–0.062*	–0.070*
	(0.015)	(0.025)	(0.035)

*Significant at the 5 percent level.
Note: All regressions include year dummy variables. Standard errors are in parentheses.
Source: Authors' calculations.

plants is statistically significant in three industries—plastics, electrical machinery, and transport equipment. Again, this result is consistent with factors that lead to divergent productivity paths for exporting and nonexporting plants, of which learning by exporting is one.

The data for Korea produce a similar pattern, but most of the differentials are not statistically significant. The coefficients in the second column indicate that exiting plants are significantly less productive than continuing exporters in only two of the five industries—textiles and electrical machinery, where the average productivity differential is 8.3 and 9.0 percent, respectively. The widening of the differential continues following exit for all but one industry, as shown in the third column, but this effect is not statistically significant in any of the industries.

Overall, the statistical evidence is stronger for Taiwan (China) than for Korea that exiting plants are less productive than continuing exporters. The evidence is also stronger for Taiwan (China) that the relative position of exporting plants continues to deteriorate after exit. And there is less evidence for Korea than for

Taiwan (China) that either productivity-driven selection or productivity improvement is correlated with export experience.

VI. Productivity Differentials between Long-Term Exporters and Nonexporters

The final comparison we make considers plants that export in both years and plants that never export (groups 4 and 1 in table 1). If the act of exporting results in higher productivity, then we should observe a divergence in the average productivity of these two groups over time. The average productivity differentials in year *t* identify the productivity premium of continuous exporters (column 2 of table 7). These results largely replicate the productivity advantage of plants that remain in the export market, as identified in table 4.

The changes in differentials over time indicate that, for most industries, the productivity of continuous exporters does not improve over time relative to that of nonexporters (column 3 of table 7). In three of the industries in Taiwan (China) and all five industries in Korea, there is no significant change in the productivity

Table 7. *Average Productivity Differences between Continuous Exporters and Continuous Nonexporters*

		Exporting firm differential	
Country and industry	*Intercept*	*Initial year*	*Change in differential over time*
Korea			
Textiles	–0.134*	0.316*	0.013
	(0.011)	(0.017)	(0.025)
Apparel	0.006	0.141*	–0.017
	(0.016)	(0.032)	(0.050)
Plastics	–0.013	0.188*	–0.032
	(0.014)	(0.027)	(0.046)
Electrical machinery and electronics	0.017	0.044	0.017
	(0.019)	(0.024)	(0.035)
Transportation equipment	–0.046	0.167*	0.037
	(0.024)	(0.038)	(0.057)
Taiwan (China)			
Textiles	–0.007	0.309*	–0.094*
	(0.010)	(0.016)	(0.023)
Apparel	–0.163*	0.307*	–0.063*
	(0.013)	(0.021)	(0.030)
Plastics	0.018*	0.191*	–0.002
	(0.005)	(0.012)	(0.017)
Electrical machinery and electronics	0.007	0.145*	0.011
	(0.007)	(0.011)	(0.015)
Transportation equipment	–0.134*	0.165*	0.042
	(0.010)	(0.019)	(0.027)

*Significant at the 5 percent level.
Note: All regressions include year dummy variables. Standard errors are in parentheses.
Source: Authors' calculations.

differential over time. In the two industries in which there is a significant change in the relative productivity of the two groups—the textile and apparel industries in Taiwan (China)—the productivity advantage of the continuous exporters falls over time. Among the group of producers in operation for the two years, there is no evidence that the average productivity of the continuous exporters rises relative to that of the plants with no export experience. There are large initial differences in productivity between the two groups, but there is no evidence that the differential widens with continued export experience. These results are not consistent with a process of ongoing learning by exporting.

A possible reason why productivity differentials between continuous exporters and nonexporters do not widen over time is that both groups are made up of plants of different ages. Learning may be concentrated among young or new plants, with older plants having already fully incorporated the knowledge acquired from past experience. To determine if this is true, we divide the plants in groups 1 and 4 into those that are new in year t and those that were already operating (either in or out of the export market) in the initial census year. We examine the productivity differentials for the new plants. The results, which are not reported here, indicate that in the transportation equipment industry in Taiwan (China), the new plants that are continuous exporters have a productivity differential that widens by 8.1 percent in year $t + 1$ relative to the new plants that have never exported. This is the only industry in Taiwan (China) for which the productivity differential widens over time. Making the same comparison for Korea, we find no industries in which the export differential widens over time. Overall, with the exception of transport equipment in Taiwan (China), this comparison provides no evidence that is consistent with the learning-by-exporting hypothesis.

VII. Summary and Conclusion

The relationships between plant-level TFP and export experience are robust and simple to summarize for the five major exporting industries in Taiwan (China). On average, exporting plants have higher productivity than nonexporters. The transition patterns reflect systematic differences in productivity: average productivity is highest for continuing exporters, followed by entrants, exiting firms, and nonexporters. Plants that diversify into the export market have higher productivity prior to entry than plants that choose not to enter and, in some industries, show evidence of productivity improvements following entry. Plants that exit the export market are less productive than those that remain in the export market. In several industries the relative position of those that exit continues to deteriorate in the years following exit. Finally, there is no evidence that the productivity advantage of continuous exporters over plants that never export increases over time.

These results are consistent with self-selection of higher-productivity plants into the export market. The evidence for several industries indicates that productivity differences between exporters and nonexporters widen as export experience accumulates; however, this tendency is limited to plants that enter or exit

the export market, not continuous exporters. This widening productivity gap could reflect direct benefits from exporting, such as knowledge spillovers from buyers, or other factors that lead to positive serial correlation in the shocks to plant productivity. In the latter case the plants with positive (negative) productivity shocks will move into (out of) the export market, and their productivity will continue to diverge from the group of plants that do not make any market transitions. Given the small number of time-series observations in our data, it is impossible to disentangle these two explanations. Nonetheless, the post-entry and post-exit patterns of change in productivity are consistent with efficiency gains that accrue from the exporting process.

Although exporters are on average more productive than nonexporters in Korea and in Taiwan (China), the productivity pattern of the cohorts moving into and out of the export market differs significantly between the two economies. In general, there is less evidence of productivity-based transitions in Korea. Prior to entry, there are no significant differences between entrants and nonentrants for three of the five industries. Following entry, there is no widening of the productivity differential between these two groups in four of the industries. This pattern is also reflected on the exit side. There is no evidence that the productivity gap between plants that exit the export market and those that remain widens after exit. Finally, there is no evidence that the productivity advantage of the group of continuous exporters widens over time relative to producers that never export. Overall, these patterns do not support the learning-by-exporting hypothesis or the self-selection hypothesis.

The lack of any strong evidence of learning by exporting is consistent with the findings of Clerides, Lach, and Tybout (1998) and Bernard and Jensen (1999). Both studies approach the question in a similar way to this article, by asking if the performance (productivity) paths of exporters and nonexporters diverge following a transition from one market to the other. However, our findings are not consistent with the micro-survey evidence reported by several authors. Rhee, Ross-Larson, and Pursell (1984) find that, among Korean firms in 1965–75, a considerable amount of production engineering knowledge came from the purchasers of Korean exporters. Similarly, Keesing and Lall (1992) study five East Asian economies in 1979–80 and find that purchasers often established buying offices in the exporting countries. These offices channeled advice on quality control, design changes, and new technologies to domestic producers. Finally, Egan and Mody (1992) study U.S. imports of bicycles and footwear from East Asian countries in the mid-1980s and find that links between industrial-country buyers and developing-country suppliers acted as conduits for information about marketing and production technologies and provided access to larger industry networks.

There are several possible explanations for the difference in findings between the two groups of studies. First, learning by exporting may have been more important as a source of expertise and knowledge in the early period of expansion of the manufacturing sector, that is, for the 1960s and 1970s in Korea. By the

middle and late 1980s, the period covered by our data, much of the knowledge had been acquired and disseminated.

Second, the knowledge gained from exporting diffuses quickly across exporters and nonexporters as a result of labor mobility among firms and other interfirm contacts. Rapid diffusion would make it less likely to observe productivity differences across the groups of exporters and nonexporters examined here. In Korea, in particular, exports expanded primarily through large business groups, rather than through small and medium-size enterprises that dominated the export expansion in Taiwan (China). If the knowledge gained from exporting is transmitted quickly among members within a Korean business group, then individual producers are likely to have incorporated much of this spillover effect prior to entry, leading to less significant productivity improvements after entry. In Taiwan (China) this transmission of knowledge must occur across firms, which is likely to be slower and less complete, leaving more opportunities for individual producers to benefit from their own export experience. This explanation is consistent with our finding of some productivity improvements following entry in four Taiwanese industries.

Third, the time-series improvements in productivity that follow from export-led learning could be small and difficult to detect compared with the cross-sectional differences in TFP. Fourth, despite the fact that all of the studies rely on micro-data, the level of industry aggregation differs significantly. Although specific products may benefit from knowledge gained through exporting, they simply are too small a share of industry production to be detected in our comparisons.

Given these qualifications, our findings suggest that the movements of producers with different levels of productivity into and out of the export market more closely reflect a process of market selection among heterogeneous plants than a process of productivity improvement flowing from export market experience. More generally, our empirical findings suggest that producer productivity is correlated less strongly with export market participation in Korea than in Taiwan (China). Several factors could contribute to this difference.

A plant's long-run expected profits from exporting should guide its export decision. Productivity may serve as a less useful indicator of long-run profits in Korea than in Taiwan (China). Total factor productivity provides a summary index of several production-related factors, including the degree of capital utilization, the importance of returns to scale, and managerial efficiency. The index shows how these factors vary across producers, but it does not provide a perfect measure of long-run expected profits. Factors other than production efficiency may be important determinants of expected profitability. If these other factors differ substantially across producers, they will tend to weaken the correlation between a producer's productivity and its pattern of participation in the export market.

Heterogeneity across producers on the demand side of the market weakens the correlation between profitability and TFP. Hobday (1995) argues that there is little emphasis among Taiwanese manufacturers on brand or product differentia-

tion and little expenditure on advertising or research and development. To the extent that export products are more homogeneous in Taiwan (China) than in Korea, profit differences and export decisions in Taiwan (China) will more closely reflect differences in productive efficiency.

Several institutional factors suggest that determinants of profitability other than productivity may be more important in Korea than in Taiwan (China). Pack (1992), Levy (1991), and Levy and Kuo (1991) argue that the dense network of subcontractors and export traders in Taiwan (China) has lowered the costs of moving into and out of the export market, particularly for small firms. By contrast, the weaker network of subcontractors and traders in Korea implies substantially higher initial investment costs for the producer, which can introduce hysteresis into the export decision. The producer's prior export experience becomes an important determinant of the decision to export and can weaken the link between current productivity and exporting choice. In the 1980s both the extent of subcontracting and entry into exporting increased in Korea, suggesting that entry and exit costs decreased gradually. Investment subsidies also decreased significantly in the 1980s. However, the effects of sunk entry and exit costs as well as of investment subsidies are likely to be long term.

Several researchers, including Pack and Westphal (1986), Westphal (1990), Levy (1991), and Rodrik (1995), have documented the importance of government investment subsidies in Korea. These policies have resulted in the channeling of credit at negative interest rates to Korea's conglomerates and provided them with insurance against business risk, particularly in the export market. In this context Korean producers are less likely to base their decisions on productivity when they consider entering, continuing, or exiting the export market. Their decisions will reflect whether they have access to the necessary finance, contacts, and insurance provided by the government.

Appendix A. Description of the Data

For Taiwan (China) we use a compilation of data from the last three industrial and commercial censuses collected by the Statistical Bureau of Taiwan's Executive Yuan. The censuses cover 1981, 1986, and 1991. The Statistical Bureau collects detailed data on each of the firms operating in the manufacturing sector, which was more than 88,000 firms in 1981 and more than 100,000 firms in 1986 and in 1991. See Aw, Chen, and Roberts (1997) for a more detailed discussion of the Taiwanese data and the construction of inputs and outputs used to measure productivity.

The data for Korea come from the census of manufactures for 1983, 1988, and 1993. The censuses cover all manufacturing plants with more than five employees in each of the 23 industries defined at the two-digit standard industry level. There were approximately 39,022 plants in 1983, 59,732 in 1988, and 88,864 in 1993.

The firm or plant observations for each country not only provide complete cross-sectional coverage of the manufacturing sector but also are matched across the censuses so that analysts can follow individual producers over time and observe entry and exit patterns. The censuses for both countries provide information on the output and input variables that are necessary to measure TFP at the firm or plant level: sales, employment, book value of the capital stock, and expenditures on labor and different types of intermediate inputs. The type of data that are collected in both countries is very similar. Therefore, we discuss the variable construction for both countries at the same time, noting differences where relevant. The type of data collected in the manufacturing census in Taiwan (China) is similar to what is collected in the United States (for its use in productivity measurement, see Baily, Hulten, and Campbell 1992) or in the industrial countries analyzed in Roberts and Tybout (1996).

For Taiwan (China) firm output is defined as total firm sales deflated by a wholesale price index defined at the two-digit industry level. For Korea the value of plant output is measured as the sum of total revenue from sales, repairing and fixing services, and subcontracted work, and the change in the inventory of final goods. It is deflated by a producer price index defined at the two-digit industry level.

We model each producer as using four inputs in production: labor, capital, materials, and subcontracting services. The labor input is measured as the number of production and nonproduction workers. Total payments to labor are measured as total salaries to both groups. The measure of capital input is the book value of the capital stock of the firm or plant. We have adjusted the book values to control for price changes in new capital goods that will cause the book values to change over time with investment in new equipment. The expenditure share on capital is calculated as the residual after subtracting expenditures on labor, material inputs, and subcontracting from the value of output.

The material input is defined to include raw materials, fuel, and electricity. In Taiwan (China) raw material expenditures are deflated by a general producer price index, which covers both manufacturing and nonmanufacturing output in the country. Fuel and electricity expenditures are deflated by an aggregate energy price index. In Korea we use a raw material price index for the manufacturing sector to deflate material expenditures. Fuel expenditures are deflated by an energy producer price index, and electricity expenditures are deflated by an electricity price index.

The final input measures expenditures on subcontracting services. Many producers in both economies hire subcontractors to perform parts of the manufacturing process. Payments to these subcontractors are reported as separate expenditures by the firm or plant in the census data. To construct the subcontracting input, we deflate payments to subcontractors by the output price of the industry in which the firm or plant operates. This is not an ideal price index for deflating subcontracting expenditures. However, the overall inclusion of the subcontracting input is important because it recognizes that the inputs of producers that

subcontract some of the production steps need to be increased, and thus their TFP reduced, relative to the producers that do not subcontract.

Appendix B. The Measurement of Plant-Level Total Factor Productivity

Using manufacturing data for Korea and Taiwan (China), we construct an index of TFP for each plant in each year. (See Tybout 1996 for a discussion of alternative productivity measures based on econometric estimation of production functions and a summary of the literature on the sources of productivity differences across producers.) In the case of Taiwan (China) this is done for each of the three census years 1981, 1986, and 1991. For Korea the three census years are 1983, 1988, and 1993.

Caves, Christensen, and Diewert (1982) develop a multilateral index that is useful for measuring TFP in plant- or firm-level panel data sets. They construct the TFP index as the log of the plant's outputs minus a revenue-share-weighted sum of the log of the plant's inputs. In order to guarantee that comparisons between any two plant-year observations are transitive, each plant's inputs and outputs are expressed as deviations from a single reference point. Caves, Christensen, and Diewert's multilateral index uses as the reference point a hypothetical plant with input revenue shares that equal the arithmetic mean of revenue shares over all observations, and output and input levels that equal the geometric mean of output and the inputs over all observations. Each plant's output, inputs, and productivity in each year are measured relative to this hypothetical plant.

Good, Nadiri, and Sickles (1997) discuss an extension of the multilateral index that uses a separate hypothetical-plant reference point for each cross section of observations and then chain-links the reference points together over time in the same way as the conventional Tornqvist index of productivity growth. This productivity index is useful in our application because it provides a consistent way of summarizing the cross-sectional distribution of plant TFP, using only information that is specific to that time period, and describing how the distribution moves over time.

Let each plant f produce a single output Y_{ft} using the set of inputs X_{ift} where $i = 1, 2, \ldots n$. The TFP index for plant f in year t is defined as:

$$\text{(B-1)} \qquad \ln TFP_{ft} = \left(\ln Y_{ft} - \overline{\ln Y_t}\right) + \sum_{s=2}^{t}\left(\overline{\ln Y_s} - \overline{\ln Y_{s-1}}\right) - \left[\sum_{i=1}^{n}\frac{1}{2}\left(S_{ift} + \overline{S_{it}}\right)\left(\ln X_{ift} - \overline{\ln X_{it}}\right) + \sum_{s=2}^{t}\sum_{i=1}^{n}\frac{1}{2}\left(\overline{S_{is}} + \overline{S_{is-1}}\right)\left(\overline{\ln X_{is}} - \overline{\ln X_{is-1}}\right)\right].$$

The first line in equation B-1 measures plant output and consists of two parts. The first part expresses the plant's output in year t as a deviation from the reference point, the geometric mean output over all plants in year t. It thus captures information on the cross-sectional distribution in output. The second part sums the change in the output reference point across all years, effectively capturing information on the shift of the output distribution over time by chain-linking the movement in the reference point. Subscript s denotes the reference year. The remaining two lines of the formula perform the same operation for each input X_i. The inputs are then summed using a combination of plant factor shares S_{ift} and average factor shares S_{it} in each year as weights. The index provides a measure of the proportional difference in TFP for plant f in year t relative to the hypothetical plant in the base year. In our application we use 1981 as the base year for Taiwan (China) and 1983 as the base year for Korea.

REFERENCES

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Aw, Bee Yan, and Geeta Batra. 1998. "Technology, Exports, and Firm Efficiency in Taiwanese Manufacturing." *Economics of Innovation and New Technology* 7(1):93–113.

———. 1999. "Wages, Firm Size, and Wage Inequality: How Much Do Exports Matter?" In David B. Audretsch and Roy Thurik, eds., *Innovation, Industry Evolution, and Employment*. Cambridge, U.K.: Cambridge University Press.

Aw, Bee Yan, Xiaomin Chen, and Mark J. Roberts. 1997. "Firm-Level Evidence on Productivity Differentials, Turnover, and Exports in Taiwanese Manufacturing." NBER Working Paper 6235. National Bureau of Economic Research, Cambridge, Mass. Processed.

Aw, Bee Yan, and Amy Hwang. 1995. "Productivity and the Export Market: A Firm-Level Analysis." *Journal of Development Economics* 47(2):313–32.

Baily, Martin Neil, Charles Hulten, and David Campbell. 1992. "Productivity Dynamics in Manufacturing Plants." *Brookings Papers on Economic Activity: Microeconomics 1992*. Brookings Institution, Washington, D.C.

Bernard, Andrew B., and J. Bradford Jensen. 1995. "Exporters, Jobs, and Wages in U.S. Manufacturing, 1976–1987." *Brookings Papers on Economic Activity: Microeconomics 1995*. Brookings Institution, Washington, D.C.

———. 1999. "Exceptional Exporter Performance: Cause, Effect, or Both?" *Journal of International Economics* 47(1):1–25.

Campa, José. 1998. "Exchange Rates and Trade: How Important Is Hysteresis in Trade?" Working Paper. Stern School of Business, New York University. Processed.

Caves, Douglas W., Laurits Christensen, and Erwin Diewert. 1982. "Output, Input, and Productivity Using Superlative Index Numbers." *Economic Journal* 92(1):73–96.

Chen, Tain-Jy, and De-Piao Tang. 1987. "Comparing Technical Efficiency between Import-Substituting and Export-Oriented Foreign Firms in a Developing Country." *Journal of Development Economics* 26(2):277–89.

Clerides, Sofronis, Saul Lach, and James Tybout. 1998. "Is 'Learning-by-Exporting' Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco." *Quarterly Journal of Economics* 113(August):903–47.

Egan, M. L., and Ashoka Mody. 1992. "Buyer-Seller Links in Export Development." *World Development* 20(3):321–34.

Ericson, Richard, and Ariel Pakes. 1995. "Markov-Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies* 62(1):53–82.

Evenson, Robert E., and Larry E. Westphal. 1995. "Technological Change and Technology Strategy." In Jere Behrman and T. N. Srinivasan, eds., *Handbook of Development Economics*. Vol. 3A. Amsterdam: North-Holland.

Good, David H., M. Ishaq Nadiri, and Robin Sickles. 1997. "Index Number and Factor Demand Approaches to the Estimation of Productivity." In Hashem Pesaran and Peter Schmidt, eds., *Handbook of Applied Econometrics. Vol. 2: Microeconometrics*. London: Basil Blackwell.

Grossman, Gene, and Elhanan Helpman. 1991. *Innovation and Growth in the World Economy*. Cambridge, Mass.: MIT Press.

Haddad, Mona. 1993. "How Trade Liberalization Affected Productivity in Morocco." Policy Research Working Paper 1096. Development Research Group, World Bank, Washington, D.C. Processed.

Handoussa, James, Mieko Nishimizu, and John Page. 1986. "Productivity Change in Egyptian Public Sector Industries after the 'Opening.'" *Journal of Development Economics* 20(1):53–74.

Hobday, Michael. 1995. *Innovation in East Asia: The Challenge to Japan*. U.K.: Edward Elgar Publishing.

Hopenhayn, Hugo. 1992. "Entry, Exit, and Firm Dynamics in Long-Run Equilibrium." *Econometrica* 60(5):1127–50.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica* 50(3): 649–70.

Keesing, Donald B., and Sanjaya Lall. 1992. "Marketing Manufactured Exports from Developing Countries: Learning Sequences and Public Support." In Gerald K. Helleiner, ed., *Trade Policy, Industrialization, and Development: New Perspectives*. Oxford: Clarendon Press.

Lambson, Val E. 1991. "Industry Evolution with Sunk Costs and Uncertain Market Conditions." *International Journal of Industrial Organization* 9(2):171–96.

Levy, Brian. 1991. "Transactions Costs, the Size of Firms, and Industrial Policy: Lessons from a Comparative Case Study of the Footwear Industry in Korea and Taiwan." *Journal of Development Economics* 34(1):151–78.

Levy, Brian, and Wen-Jeng Kuo. 1991. "The Strategic Orientation of Firms and the Performance of Korea and Taiwan in Frontier Industries." *World Development* 19(4):363–74.

Pack, Howard. 1992. "New Perspectives on Industrial Growth in Taiwan." In Gustav Ranis, ed., *Taiwan from Developing to Mature Economy*. Boulder, Colo.: Westview Press.

Pack, Howard, and Larry Westphal. 1986. "Industrial Strategy and Technical Change: Theory versus Reality." *Journal of Development Economics* 22(1):87–128.

Rhee, Yung W., Bruce Ross-Larson, and Garry Pursell. 1984. *Korea's Competitive Edge: Managing the Entry into World Markets*. Baltimore, Md.: Johns Hopkins University Press.

Richardson, J. David, and Karin Rindal. 1995. *Why Exports Really Matter!* Washington, D.C.: Institute for International Economics and the Manufacturing Institute.

———. 1996. *Why Exports Matter: More!* Washington, D.C.: Institute for International Economics and the Manufacturing Institute.

Roberts, Mark J., and James R. Tybout. 1997. "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs." *American Economic Review* 87(September):545–64.

Roberts, Mark J., and James R. Tybout, eds. 1996. *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure.* Oxford: Oxford University Press.

Rodrik, Dani. 1995. "Getting Interventions Right: How South Korea and Taiwan Grew Rich." *Economic Policy* 20(1):55–107.

Tybout, James R. 1996. "Heterogeneity and Productivity Growth: Assessing the Evidence." In Mark J. Roberts and James R. Tybout, eds., *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure.* New York: Oxford University Press.

Tybout, James R., and M. Daniel Westbrook. 1995. "Trade Liberalization and Dimensions of Efficiency Change in Mexican Manufacturing Industries." *Journal of International Economics* 31(August):53–78.

Westphal, Larry. 1990. "Industrial Policy in an Export-Propelled Economy: Lessons from South Korea's Experience." *Journal of Economic Perspectives* 4(1):41–59.

World Bank. 1993. *The East Asian Miracle: Economic Growth and Public Policy.* New York: Oxford University Press for the World Bank.

THE WORLD BANK ECONOMIC REVIEW, VOL. 14, NO. 1: 91-109

Vintage Technologies and Skill Constraints: Evidence from U.S. Exports of New and Used Machines

Giorgio Barba Navaretti, Isidro Soloaga, and Wendy Takacs

When countries import production machinery, they must choose between new and used equipment. This article looks at that choice in the presence of labor-saving technical progress and complementarity between technologies and skills within the firm. It develops a theoretical model of the market for used machines. It then analyzes data on U.S. exports of metalworking machine tools by country of destination, classifying machines according to their vintage and their technological characteristics. The data show that the share of used equipment imported is higher if the importing country has a lower level of development, as measured by income per capita. Econometric estimation of the determinants of this share shows that it also is higher the greater is the technological change embodied in the machine or the greater is the change in the skills required to run the machine efficiently.

These results indicate that technological factors and skill constraints may be as important as factor prices in determining the choice of technique in developing countries. The policy recommendation emerging from this work—avoid constraints on imports of used equipment—is similar to that in the existing literature. But the reasoning is different. Instead of emphasizing inappropriate capital-labor ratios for low-wage countries, the results indicate that investment in advanced technologies is effective only if importing countries have the skills to use them.

Many developing countries design their trade policies to discriminate against importation of secondhand goods, imposing import bans, licensing requirements, or higher tariff rates. Some industrial countries even discriminate against used products: witness Australia's additional $12,000 tariff on used cars (Wonnacott

Giorgio Barba Navaretti is with the Department of Economics at the Universitá degli Studi di Ancona, Centro Studi Luca d'Agliano, and Fondazione Eni Enrico Mattei; Isidro Soloaga is with the Development Research Group at the World Bank and the Department of Economics at the University of Maryland, College Park; and Wendy Takacs is with the Department of Economics at the University of Maryland, Baltimore County. Their e-mail addresses are *barba@mailserver.unimi.it, isoloaga@worldbank.org,* and *wtakacs@umbc.edu.* This article was prepared while Mr. Barba Navaretti was a visiting researcher with the Development Research Group at the World Bank. Fondazione Eni Enrico Mattei and Centro Studi Luca d'Agliano partly funded this project. Mr. Anthony M. Bratkovich, engineering director of the U.S. Association for Manufacturing Technology, was extremely patient and helpful in leading the authors through the mysteries of the technological characteristics of metalworking machines. The authors appreciate the useful suggestions of Gabriel Castillo, John Nash, Garry Pursell, Maurice Schiff, James Tybout, and the anonymous referees. They also thank Jeff Hayden and Jennifer Ngaine for their assistance.

1994). These policies are motivated by a desire to protect domestic industries against competition from low-priced goods, to avoid becoming a dumping ground for castoffs from high-income countries, to push industries toward the technological frontier, and to avoid the use of obsolete technologies.

But trade restrictions on used capital goods appear contrary to the appropriate choice of production techniques in developing countries, where low wages and high interest rates call for the use of labor-intensive production processes. Older equipment is likely to be more labor-intensive than new equipment; technological change tends to be labor-saving, and older equipment requires greater maintenance and carries a greater risk of downtime. Moreover, the optimal scale of older machines is smaller, which may be more appropriate for the smaller markets in developing countries, and older machines may be more flexible in their use and less specialized. Thus several authors conclude that firms in low-wage developing countries would find secondhand equipment more profitable than new machines and that developing countries would suffer a welfare loss from import restrictions on used machinery (Sen 1962, Schwartz 1973, James 1974, Thoumi 1975, Mainwaring 1986, and Pack 1977).

In this literature some models focus on the impact of greater maintenance costs as machines age (Schwartz 1973 and Thoumi 1975). The literature on vintage capital emphasizes labor-saving technological change (Bardhan 1970, Smith 1976, Gabisch 1975, and Pack and Todaro 1970). And some models incorporate both phenomena (Mainwaring 1986). Recent contributions on technology transfer link the choice of technique to the skills available in a firm or country. These skills are human capital or other technological capabilities acquired through deliberate learning or learning by doing (Benhabib and Rustichini 1991, Chari and Hopenhayn 1991, Keller 1994, Jovanovic and MacDonald 1993, and Jovanovic and Nyarko 1995, 1998). The more skills that are specific to a particular technique, the more costly it is to switch to that technique. The skill factor is likely to affect the choice between new and used machines when new machines embody technical change.

In this article we model a firm's choice between new and used machines. We test the predictions emerging from the modeling exercise using data on U.S. exports of new and used metalworking machinery, disaggregated by type of machine and by country of destination. The model incorporates three factors: greater downtime as machines age, labor-saving technical progress, and the greater skill requirements of more technologically sophisticated machinery.

Most of the literature on trade in used machinery focuses on heterogeneity between countries based on the stylized fact that developing countries have lower wages and higher capital costs than industrial countries (Sen 1962, Smith 1974, and Mainwaring 1986). We adopt a slightly different model based on heterogeneity among firms (as in Bond 1983). Our model takes into account that if labor and capital markets are imperfect or distorted by sectoral labor regulations or subsidized directed credit, firms in the same country may face different wage rates and capital costs.

Firms also can differ in the technical and managerial skills available to them. Heterogeneity among firms located in different countries provides the underlying motive for international trade in new and used capital equipment. Models that do not take firm heterogeneity into account predict fairly extreme patterns of trade in used machinery. For example, several models predict that firms in developing countries would import only the oldest machinery available. The assumption that firms within developing countries may face different wage structures and interest rates is reasonable, given imperfections in capital markets, the coexistence of multinationals and purely domestic firms, and the dichotomy between the formal and informal sectors.

I. The Firm's Choice between New and Used Machinery

This section develops a model that is a version of Bond (1983), modified to include international trade of machinery.

Differences between New and Used Machinery

New and used machinery can differ in three important ways: risk of breakdown, productivity of embodied technology, and required technical skills. Used machinery normally requires more maintenance and is more likely to break down than new machinery. Maintenance demands high labor input. In addition, if employees are paid for a regular work schedule, machine downtime means that workers are idle, implicitly increasing the labor intensity of the production process. We capture the impact of breakdowns and the maintenance requirements of used machines by adjusting output for downtime using the factor α, defined as the ratio of a used machine's output to that of an identical new machine, that is, one not yet operated ($0 < \alpha < 1$).

Machinery of a given vintage embodies the technology available when it was produced. Labor-saving technical progress is captured by distinguishing between output per worker with a machine newly produced in the current period (a_n) and output per worker with a machine embodying last period's technology, when that machine was new (a_u). Thus the ratio a_u / a_n captures the rate of labor-saving technical progress, independent of the downtime effect, α, with $(a_u / a_n) \leq 1$.

To identify the independent impacts of technical progress and the aging process in the context of metalworking machines, we distinguish between low-tech and high-tech machines. From a technological point of view machine tools (especially metal-cutting tools) can be divided into two broad categories. Numerically controlled machines have a high rate of technological upgrading, linked to the development of electronics. These are "high-tech" machines. Manual machines may improve in terms of design and safety, but they have no or a very low rate of technical progress. These are "low-tech" machines. The difference in labor productivity between new and used low-tech machines is attributable only to increases in maintenance and longer downtime. The difference in labor produc-

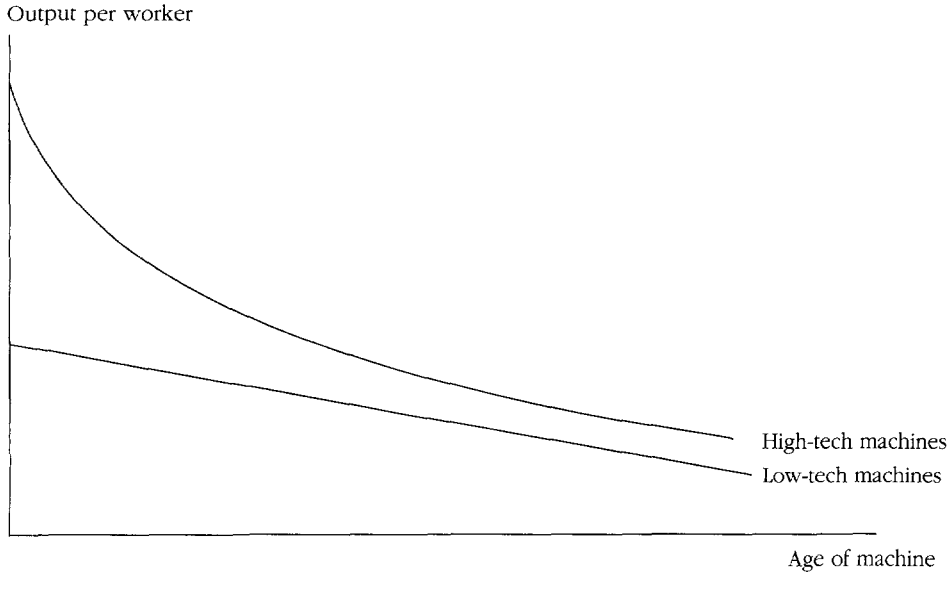
tivity between new and used high-tech machines is attributable to technical progress as well as to increases in maintenance and longer downtime.

For low-tech (manual) machines output per worker varies with age because of the downtime (α) effect (figure 1). For high-tech (numerically controlled) machines labor productivity declines with the age of the machine because of the combined effect of downtime and the lower technological sophistication embodied in older machines (figure 1). There are two key features of the relationships between labor productivity and the two types of machines. First, output per worker is always lower for low-tech than for high-tech machines. Second, at any point in time, the decline in labor productivity with age is larger for high-tech machines than for low-tech machines.

The third way in which new and used technologies may differ is in the skills required to operate the machines. The literature on vintage technology emphasizes the role of technology-specific skills (Evenson and Westphal 1994 and Keller 1994). Metalworking machine tools provide a good example. Manual machines (low-tech) require sophisticated craftsmen to operate them. Numerically controlled (high-tech) machines require electronic technicians. Accumulated learning by doing could be lost when a firm switches to a new technology (Chari and Hopenhayn 1991, Dasgupta and Stiglitz 1988, Jovanovic and MacDonald 1993, and Jovanovic and Nyarko 1995, 1996). Complementarities between workers with different skills may constrain the choice of technology (Chari and Hopenhayn 1991).

Linking the educational level of the people using the machines (craftspeople, technicians, engineers) to the technology embodied in the machines is essential,

Figure 1. *Labor Productivity and Machine Age with and without Technical Progress*



but not sufficient. Technological knowledge is often tacit and not transmittable in codified form (David 1993), and technological capabilities are related to the performance of many different technological functions (Lall 1987). Skills, then, should refer to the absorptive capacity of a firm or a country, that is, the ability to master a given technology (Evenson and Westphal 1994 and Keller 1994). Absorptive capacity is affected by the physical, social, and economic characteristics of a firm or country.

Firms therefore may be reluctant to move to high-tech machines because they do not have the skills to use them or because building up such skills would be more costly than continuing to use low-tech machines. In our simplified setting we assume that if a firm adopts a new high-tech machine that it does not have the appropriate skills to run it suffers a loss in productivity. The firm's current level of productivity is captured by γ, the proportion of full-capacity output achievable with new machines, given current skills (where $0 \leq \gamma \leq 1$). We refer to γ as the "inability coefficient," since γ is lower the less able is the firm. The skill factor may constrain the choice between new and used machines to the extent that new machines embody an increasing level of technological sophistication.

Trade Policy

Different trade policy instruments influence the choice between new and used machines in different ways. An equal ad valorem tariff rate, t, on all imported machines raises the domestic price of new and used machines proportionally. Restrictions discriminating against used equipment through either higher tariffs or licensing requirements increase the cost of used equipment disproportionately, discouraging its use.

Relative Prices

We assume that machines last for two periods. A firm buying a new machine can sell it at the end of the period for the going price of a used machine, but used machines will have zero scrap value at the end of the second period of use. The analysis assumes that machines are paid for at the beginning of the period and that wages are paid out at the end of the period. At the end of the period the firm's (net) cost of using a new machine embodying the current period's technology would be:

$$C_n = P_n(1+r)(1+t) - P_u(1+t) + \frac{w}{a_n}q. \tag{1}$$

The firm's (net) cost of using a machine embodying the previous period's technology (if the machine had never been used) would be:

$$C_u = P_u(1+r)(1+t) + \frac{w}{a_u}q \tag{2}$$

where C_i is the total cost of production using machinery i ($i = n$ for new machines, and $i = u$ for used machines), P_i is the price of machine i, a_i is the labor productivity of machine i ($a_i = q / L_i$) when it is (was) unused, L_i is the labor input per time period with machine i, w is the wage rate per time period, r is the interest rate, t is the ad valorem tariff rate on machinery imports, and q is the full-capacity output of a machine when it is new.[1]

If the machine embodying last period's technology is used in that period, it yields only α of the output it did when new. If the productivity of the new machine is constrained because of the lack of skill in the labor force, the new machine yields only γ times the designed output.[2] More precisely, α is the proportion of the full-capacity output of a new machine that can be produced when the machine has been used for one period, and γ is the proportion of full-capacity output that can be produced using the new machine with current skills.

A firm will be indifferent between new and used machinery if unit costs are the same with the two types of equipment:

$$\frac{C_n}{\gamma q} = \frac{C_u}{\alpha q}. \tag{3}$$

Thus a firm will be indifferent between new and used equipment when:

$$\frac{P_n(1+r) - P_u(1+t) + \frac{w}{a_n}q}{\gamma} = \frac{P_u(1+r)(1+t) + \frac{w}{a_u}q}{\alpha}. \tag{4}$$

Solving for P_u yields U, the price of used equipment at which the firm is indifferent between using new and used equipment:

$$U = \frac{\alpha P_n - \gamma\beta\left(1 - \frac{\alpha a_u}{\gamma a_n}\right)\frac{wq\theta}{a_u}}{\gamma + \alpha\beta}, \tag{5}$$

where $\beta = 1 / (1 + r)$ and $\theta = 1 / (1 + t)$.

If the market price of used machinery (P_u) is less than a firm's indifference price (U), the firm will buy used equipment; if the market price is greater ($P_u > U$), the firm will buy new equipment. Given the market price, P_u, an increase in U makes it more likely that a firm will buy used equipment, and a decrease in U raises the chance that it will buy new equipment.

1. We assume that full-capacity output of a new machine is always the same (q) independent of the type of machine, whereas the labor input necessary to achieve full-capacity output changes (thus affecting labor productivity a_i).

2. Labor input does not change if the machine is used at less than full-capacity output. Either workers are paid on a fixed schedule, or they use their idle time to maintain machines.

Equation 5 implies the following:

- The indifference price of used machines equals the price of new machines with production capacity equal to that of used machines (the first term in the numerator on the right side of equation 5), net of the higher labor costs of used compared with new machines (the second term in the numerator on the right side of equation 5).
- Other things being equal, the smaller is α (the more the use of equipment generates a loss in productivity) and the smaller is a_u / a_n (the greater is the rate of labor-saving technical progress), the lower is the indifference price (the less desirable are used machines).
- The indifference price of used machines increases as γ falls, other things being equal, so firms that do not have the technical skills required to run new machinery will be more likely to opt for used equipment.

The impact of the wage rate, interest rate, and machinery tariff rate on the indifference price of used machinery is more complicated to assess and depends crucially on the firm's skill level, γ. Unless the firm lacks most of the skills needed to use higher-technology equipment, the productivity of new machines is greater than that of used machines for the firm, that is, $\alpha a_u / \gamma a_n < 1$. In this case U declines when w and β increase, so firms facing high wages and low capital costs are more likely to prefer new machines (because it is more likely that $U < P_u$). Also, U decreases as θ increases, implying that higher tariffs raise the indifference price of used machines, making their purchase more likely. Indeed, tariff increases push the indifference price downward in the same way as do increases in the cost of capital (r). But lack of skills to make use of higher-technology equipment (a sufficiently low γ) could eliminate or reverse the influence of wages, interest rates, and tariff levels on the indifference price.

The Market for Used Machines

We assume that there are two regions, North and South. Firms in the North (N) are homogeneous (all have the same γ) and face the same factor prices. Firms in the South are either large (L) or small (S). Large southern firms have more technical skills (higher γ) and face more expensive labor and cheaper capital than small southern firms.[3]

Given the homogeneity of northern firms, used machines are supplied at their indifference price (U_N), therefore, $P_u = U_N$. Northern firms are indifferent between purchasing used or new machines.[4] The South is a price taker because it is

3. This assumption is consistent with widespread evidence that labor and capital markets in developing countries are fraught with imperfections and often are segmented, so that small and large firms face different wage and interest rates. These differences are a result of credit rationing, labor regulations, dichotomies between formal and informal sectors, and the coexistence of multinational and indigenous firms. In addition, small and large firms differ in terms of technical skills and therefore may have widely different values of γ.

4. The northern indifference price is the equilibrium price of used machines even if northern firms only use new machines.

small in machinery markets compared with the North. New and used machinery prices are determined in northern machinery markets, and both large and small southern firms face an infinitely elastic supply of new and used machines at those prices. In the South large firms have sufficient skills to absorb new technology (high γ). They face sufficiently high wages and low interest rates so that, at the prices of new and used machines determined in the North, large southern firms would opt for new machines, as $U_L < U_N$. Small southern firms, in contrast, have relatively low γ and face high interest rates and low wages. Thus the indifference price of small southern firms exceeds that of northern firms (from equation 5), and small firms buy only used machines. With heterogeneous firms in the South and homogeneous firms buying at the indifference price in the North, firms in both the South and the North purchase new and used machines.

In the South small firms' demand for used machines and large firms' demand for new machines depend on the demand for their output. Suppose that small and large firms' products (X_S and X_L, respectively) are imperfect substitutes, are nontradable, and have downward-sloping demand functions $P(X_i)$. The demand for used machines by small southern firms must be consistent with the zero-profit condition:

$$P_u = \left[\alpha P(X_S) - \frac{w_S}{a_u}\right] q\beta_S\theta, \tag{6}$$

where X_S is the quantity of the final product produced and sold by small southern firms, $P(X_S)$ is the price of X_S, and subscript S designates the value of the variables for small southern firms.

The production function for small southern firms is given by:

$$X_S = \alpha q Q_u, \tag{7}$$

where Q_u is the quantity of used machines employed by small southern firms.

The demand function for good X_S is given by:

$$P(X_S) = b - cX_S. \tag{8}$$

Substituting equation 7 into equation 8, and equation 8 into equation 6, yields the demand function for used machines by small southern firms:

$$P_u = (\alpha b - w_S/a_u) q\beta_S\theta_S - c\alpha^2 q^2 \beta_S\theta_S Q_u^{DS}. \tag{9}$$

When $P_u > U_S$, that is, when the price of used equipment is greater than the indifference price of southern firms, they will not buy any used machines.[5] Thus

5. Technology and the cost parameters of large and small firms are assumed to be such that large southern firms strictly prefer new machines and small southern firms strictly prefer used machines at the northern supply price, U_N. The indifference price of small firms, U_S, is the price of used machines at which small firms are indifferent between producing and not producing. We use a somewhat different notation to distinguish this concept. We thank an anonymous referee for pointing out this subtlety.

the inverse demand function (equation 9) must be consistent with the following condition: $P_u \geq U_S$ when $Q_u = 0$.

Therefore, the demand for used machines among small southern firms is given by:

$$P_u = U_S - c\alpha^2 q^2 \beta_S \theta Q_u^{DS}. \tag{10}$$

In equilibrium, because the North is large and pins down the price for used equipment, $U_N = P_u$, and the quantity of used equipment demanded by the South will be the quantity demanded at the North's indifference price (U_N). Thus the equilibrium quantity of used machines demanded by small southern firms is given by:

$$Q_u^{DS} = \frac{U_S - U_N}{c\alpha^2 q^2 \beta_S \theta}. \tag{11}$$

Equation 11 shows that the quantity of used machines demanded in the South is greater the larger is the gap between the indifference prices of small southern and northern firms.

Similarly, the demand for new machines by large southern firms must be consistent with their zero-profit condition:

$$P_u = P_n(1 + r_L) + \frac{w_L}{a_n} q\theta - \gamma_L P(X_L) q\theta, \tag{12}$$

where subscript L designates the value of the variables for large southern firms.

Their production function is given by:

$$X_L = \gamma_L q Q_n, \tag{13}$$

where Q_n is the quantity of new machines employed by large southern firms.

The demand function for good X_L is given by:

$$P(X_L) = e - gX_L. \tag{14}$$

Substituting equation 13 into equation 14, and equation 14 into equation 12, yields large southern firms' demand for new machines as a function of the price of used machines:

$$P_u = P_n(1 + r_L) - \left(\gamma_L e - \frac{w_L}{a_n}\right) q\theta + g\gamma_l^2 q^2 \theta Q_n^{DL}. \tag{15}$$

When $P_u < U_L$, that is, when the price of used equipment is less than the indifference price of large southern firms, large southern firms will not buy new machines. Thus the inverse demand function (equation 15) must be consistent with the following condition: $P_u \leq U_L$ when $Q_n^{DL} = 0$.

Therefore,

$$P_u = U_L + g\gamma_L^2 q^2 \theta Q_n^{DL}. \tag{16}$$

Again, the equilibrium price of used equipment is the northern indifference price (U_N). Thus the equilibrium quantity of new machines demanded by large southern firms is given by:

$$Q_n^{DL} = \frac{U_N - U_L}{g\gamma_L^2 q^2 \theta}. \tag{17}$$

In the South large firms keep new machines for one year and then sell them to small firms. Small firms can buy used machines from large firms or import them. In the steady state, large firms' demand for new machines (Q_N^{DL}) will equal their supply of used machines (Q_U^{SL}), and the domestic market for used machines in each southern firm will be in equilibrium when:

$$M_u = Q_u^{DS} - Q_u^{SL}, \tag{18}$$

where M_u is the quantity of used machines imported.

The equilibrium quantity of used machines imported by the South will be determined by the interaction between the northern indifference price, U_N, which defines foreign supply (*NN* in figure 2), large southern firms' supply of used machines (equation 17 and *LL* in figure 2), and small southern firms' demand for used machines (equation 10 and *SS* in figure 2). Given equation 18, the quantity of imports of used machines is the horizontal distance between *LL* and *SS* at height *NN*.[6]

When the gap between small southern firms' indifference price and the world price of used machines increases, the shifts in *SS* and *NN* generate an increase in the quantity of used machines demanded by small southern firms. When the gap between the world price of used machines and large southern firms' indifference price increases, the shifts in *LL* and *NN* generate an increase in the domestic supply of used machines.

The ratio Φ_U of used machines imported to total machines imported (estimated empirically in the following section) will then be

$$\Phi_U = \frac{M_u}{M_u + Q_n^{DL}} = \frac{Q_u^{DS} - Q_m^{DL}}{Q_u^{DS}} = 1 - \frac{Q_n^{DL}}{Q_u^{DS}}.$$

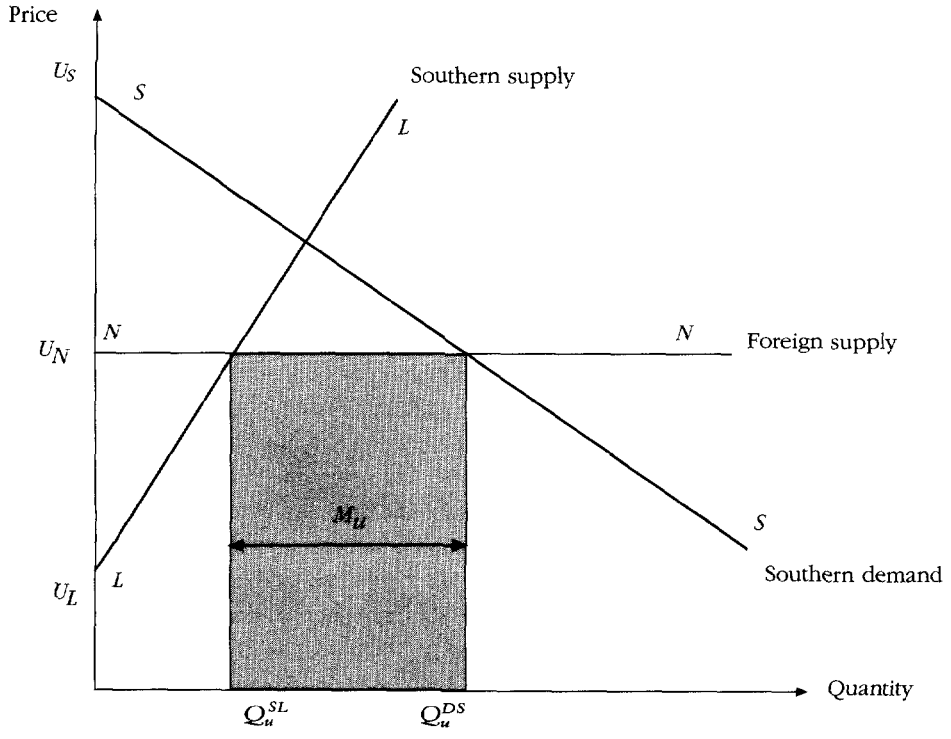
Therefore,

$$\Phi_U = 1 - \frac{(U_N - U_L)c\alpha^2\beta_S}{(U_S - U_N)g\gamma_L^2}. \tag{19}$$

Equations 5 and 19 show how the share of used equipment imported depends on the technological complexity of the machine, the technical skills of the

6. Figure 2 was drawn under the assumption that the southern country is a net importer of used machines because domestic supply is lower than total domestic demand.

Figure 2. *The Market for Used Machines in the South*



importing firm, relative prices, trade policies, and consumers' demand for final output.

First, small southern firms' total demand for used machines increases with technical progress. Indeed, the greater is the technical progress differentiating new and used equipment (the lower is a_u/a_n), the larger is the $U_S - U_N$ gap for any given difference in factor prices. An increase in technical progress lowers the indifference prices of both northern and southern firms. Both *NN* and *SS* shift downward, and the equilibrium quantity of used equipment demanded in the South increases. However, because small firms in the South have lower wages, their indifference price declines less than that of the North.

An implication of this result is that faster technical progress (smaller a_u/a_n) will increase the technological gap between northern and southern capital stocks. The increase in total demand for used machines does not necessarily result in an equal increase in imports of used machines because the increase in demand can be counterbalanced partly by the downward shift of *LL*. If wages in large southern firms are higher and interest rates are lower than those in the North, the downward shift in *LL* will be greater than the downward shift in *NN*, and the domestic supply of machines will rise. However, if the lower indifference price for large southern firms is due to greater technical skills (higher γ), *LL* will shift less than *NN*, and the domestic supply of used machines will decline.

Second, the net effect of an increase in the downtime of used machines (a decline in α) on the indifference prices is the same as the one caused by technical progress: *SS*, *LL*, and *NN* shift downward, and the net demand for used machines is expected to increase. Moreover, *SS* becomes flatter because of the effect of the decline in output per machine on the equilibrium price of output.[7]

Third, a lower γ (a less-skilled labor force) in small southern firms implies a higher southern indifference price and therefore a larger demand for used machines. Indeed, the increase in productivity of new machines will be offset by the lack of appropriate skills to use them. Equally, a lower γ in large southern firms implies an upward shift in *LL* and a reduced supply of used machines. Thus when technical skills in the South are lower, the share of used machines imported is greater.

Fourth, factor prices and tariffs affect the equilibrium quantity of used machines in the expected directions. A relative increase in northern wages lowers the northern indifference price, increases the quantity of used machines demanded by small southern firms, and reduces the domestic supply of used machines.[8] The share of used machines imported increases. Indeed, the price of used machines would decline more than southern returns from used machines. The opposite occurs when there is a relative increase in southern wages. If wages in small southern firms increase, the domestic demand for used machines declines. If wages in large southern firms increase, the domestic supply of used machines increases. For the same reason, a decline in northern interest rates raises the equilibrium quantity of used machines bought by the South.

Fifth, the choice of machines also is affected by demand for the two products X_L and X_S. If consumers' preferences shift toward one of the two products, the demand for machines will change accordingly. This effect is captured by parameters c and g in the two demand functions.

II. Determinants of Trade in Used Equipment

In this section we apply the model to data on U.S. exports. U.S. export data on some types of vehicles, equipment, and machinery are sufficiently disaggregated to distinguish between new and used goods. We concentrate on U.S. exports of metalworking machine tools in 1990–94, disaggregated by commodity classification, whether they are new or used, and country of destination. The sample covers 38 types of metalworking machines, aggregated to the six-digit level in the Harmonized System, that are exported to 23 countries.

Different types of metalworking machine tools require different types of skills. Manual machines are operated by skilled workers or craftspeople. Numerically controlled machines are operated by both technicians and unskilled workers. Machining centers are even more complex, demanding higher-level technicians

7. In equations 11 and 18, α is multiplied by c, which is the slope of the demand function of X_S.

8. This holds if $(\alpha a_u / \gamma a_n) < 1$, from equation 5.

Table 1. *The Skill Requirement Index*

Value of index	*Minimum level of worker required*
1	Unskilled labor
2	Skilled craftspeople
3	Technicians
4	Higher-level technicians, engineers

Source: Authors' calculations.

and engineers. Furthermore, machines with higher skill requirements tend to have a faster rate of technological change. Thus firms upgrading from manual to numerically controlled machines must change their endowment of skills. This requirement probably explains why only 25 percent of new investments in machine tools made in the United States between 1985 and 1989 were in numerically controlled machines, although U.S. manufacturers started investing in these types of machines in the early 1970s (Oliner 1993).

With the guidance of an engineer intimately familiar with the complexities and skill requirements of each type of machine, we developed a skill index for each 10-digit export category, reflecting the degree of skill required to operate that type of machine. The value of the index ranges from 1 to 4, increasing with the level of skill required (table 1). Looking at the shares (by quantity) of used machinery in total imports of machinery from the United States, we see, as expected, that low-income countries import a higher ratio of used to new machinery (table 2). But the variation in the shares of machines imported secondhand is not huge (between roughly 10 percent for high-income countries and about 24 percent for low-income countries). The average skill index of imported machinery is higher for high-income countries than for low-income countries, but again the difference is not large. If we divide machines into high-tech (skill indexes 3 and 4) and low-tech (skill indexes 1 and 2) categories, the same pattern emerges. The ratio of used machines to new machines imported is greater for low-income countries than for high-income countries. In addition, the share of equipment imported secondhand is larger for high-tech than for low-tech machines.

These figures provide some empirical support for our hypotheses on trade in used equipment. However, they do not tell us which of the factors cited as determining the choice between new and used machines are significant. To cast some light on this issue, we use econometric analysis.

For each category of machinery we estimate the share of used machinery in total U.S. exports to each importing country as a function of variables specific to the importing country (trade barriers and levels of education and development), variables specific to the machines (skill level required and a proxy for the rate of technical progress), and a variable combining both country and machine factors (wage-rental ratio). The basic estimating equation is:

$$Q_{ij}^{u} = \alpha_0 + \alpha_1 W_{ij} + \alpha_2 T_{ij} + \alpha_3 S_{ij} + \alpha_4 E_j + \alpha_5 t_{ij} + \alpha_6 Y_j + \nu_{ij}. \tag{20}$$

Table 2. *Imports of Metalworking Machine Tools from the United States*

Importing countries [a]	*Ratio of used to new machinery imported*	*Average index*[b]	*Ratio of used to new machinery imported*	
			Low skill (skill indexes 1 and 2)	*High skill (skill indexes 3 and 4)*
High-income	0.096	2.93	0.082	0.296
Middle-income	0.112	2.71	0.095	0.339
Low-income	0.235	2.63	0.159	0.526

a. High-income countries have GDP per capita greater than $12,000, middle-income countries have GDP per capita between $1,300 and $12,000, and low-income countries have GDP per capita less than $1,300.
b. Weighted average by value of shipments.
Source: Authors' calculations.

where Q^u_{ij} is the quantity of used machinery of type i exported to country j as a proportion of total machinery of type i imported by country j between 1990 and 1994; W_{ij} is the wage-rental ratio, which is log $\{w_j / [P_i (i_j + d)]\}$, where w_j is the average annual wage in dollars (1990–94), P_i is the unit value of new machinery of type i, i_j is the average real interest rate in country j, and d is the depreciation rate (assumed to equal 10 percent);[9] Y_j is gross domestic product (GDP) per capita in dollars in country j (1990–96 average); T_{ij} is the percentage difference between the unit values of new and used machinery of type i imported by country j, which we use as a proxy for the rate of technological change; S_{ij} is the average skill requirement index for machines of type i exported to country j; E_j is the education level in country j, measured by average years of schooling (over 1990–96); t_{ij} is the tariff rate on imports of machinery of type i in country j; and v_{ij} is the disturbance term. See the appendix for data sources.

Consistent trade data for new and used machinery are available only as far back as 1990.[10] Many country-specific variables are not available for all of these years, so we aggregate the five years. We use total U.S. exports of each commodity classification to each country in 1990–94 for trade flows and average observations on the other variables across the five years or over the largest number of years for which we have observations.

We estimate two versions of the model, including and excluding per capita GDP as a general indicator of development (table 3). Initial estimates using the ordinary least squares (OLS) estimation method (which are not reported) posed heteroskedasticity problems because the variance of the error term is decreasing with the share of imports of machines of type i in total imports of country j. Indeed, when the share is lower, imports of a given machine are less stable. We

9. The results were not sensitive to experiments using depreciation rates of 7.5 and 12.5 percent. In earlier versions of this article we attempted to use wages or ratios of wage to interest rates rather than the more appropriate ratio of wage to rental rates. The estimated coefficients of these variables were uniformly insignificant at standard confidence levels.

10. A reclassification of the export data in 1990 created a break in the series, so although data on used machinery exports are available for previous years, they are not based on exactly the same commodity classifications.

Table 3. *Determinants of Imports of Used Machinery*

Variable	*1a* Weighted OLS	*1b* White corrected	*2a* Weighted OLS	*2b* White corrected
Constant	0.602***	0.467***	0.504343***	0.491***
	(5.047)	(3.507)	(3.048)	(2.810)
Wage-rental ratio	–0.024*	–0.026**	–0.02782**	–0.025*
	(–1.928)	(–2.078)	(–2.103)	(–1.947)
Technical change	0.066***	0.057***	0.064734***	0.058***
	(5.803)	(4.985)	(5.638)	(4.938)
Skill requirement	0.051**	0.047*	0.046061*	0.048*
	(2.206)	(1.795)	(1.796)	(1.813)
Education	–0.043***	–0.013	–0.047299***	–0.011
	(–3.314)	(–0.975)	(–3.379)	(–0.723)
GDP per capita			0.021825	–0.005
			(0.857)	(–0.212)
Tariff	–0.003*	–0.0001	–0.003028*	–0.001
	(–1.736)	(–0.632)	(–1.751)	(–0.627)
Adjusted R^2	0.3974	0.3464	0.3967	0.3436
F-value	30.941	25.066	25.876	20.804
Number of observations	228	228	228	228

*Significant at the 10 percent level.
**Significant at the 5 percent level.
***Significant at the 1 percent level.
Note: The dependent variable is the share of machinery imported secondhand. *t*-statistics are in parentheses. Data are for imports of U.S. metalworking machines for 1990–94.
Source: Authors' calculations.

correct for heteroskedasticity using two methods. The first is a weighted least squares method, using as weights the share of imports of machines of type *i* in total imports of country *j*. Because this method may generate endogeneity problems that risk introducing spurious correlation, we also use White's method of correcting standard errors in the presence of heteroskedasticity.

As expected, the results for all of the equations indicate that factor costs affect the choice between new and used capital equipment. An increase in the wage-rental ratio shifts imports toward new and away from used machines.

The results also indicate the importance of technological and skill factors in the choice between new and used equipment. The positive and significant estimated coefficient of the machine-specific skill index indicates that the more high-tech are the machines, the greater is the proportion of equipment imported secondhand. The performance of the machine-specific skill variable is strong and robust across all versions of the estimated equations. The results also indicate that the faster is the rate of technical progress (as measured by our proxy, the percentage drop between new and used equipment prices), the larger is the share of capital equipment imported secondhand.

The rationale behind using the price differential between new and used machines as a proxy for the rate of technical progress is that greater improvement in a new machine results in a greater drop in the market price of the previous model. We recognize, however, that this measure is an imperfect proxy because the gap

between the prices of new and used machines could be influenced by other factors, such as asymmetric information (the "lemon" phenomenon). In addition, the measured technology variable could include variation in the average price of a given type of machinery across countries as well as variation across types of machines. Machines in the used category could range from barely used to barely usable. For countries that choose very old (and therefore much cheaper) machines, the relative price of new machines will be high. Thus the estimated coefficient may be biased because the measured technology variable could be correlated with the factors that make the country more prone to select used machines.[11] Bearing in mind these caveats, the evidence that technological factors are important in the choice of new or used machinery should be interpreted with some caution.

The estimated coefficient of the country-specific, skill-related education variable has the expected sign and is significant in the weighted OLS estimations, but not in the equations estimated using White's corrected standard errors. The education variable was measured as average years of schooling, which may be too general to capture the extent of technical training. We include GDP per capita as another potential measure of a country's technological development, but the results are insignificant.[12] The results for the tariff variable are somewhat disappointing. The model predicts that tariffs on machinery, by increasing the cost of capital to the firm, would encourage firms to opt for more labor-intensive used machinery. Yet the estimated coefficient of the tariff variable is negative and at times significant.[13]

III. Conclusions

Developing and transition economies frequently discriminate against imports of secondhand goods, including production machinery. The literature on this issue points out that these restrictions are costly because they deny firms access to older equipment, which is usually more labor intensive than new equipment and thus more appropriate for low-wage countries. In this article we developed a model that extended the established approach to take into account technological

11. We thank an anonymous referee for this point. This bias could possibly be corrected by using U.S. domestic data on the new-to-used price differential across different types of machines to construct the proxy for the technical change variable. Unfortunately, even if these data were readily available, difficulties and uncertainties in constructing a cross-classification between domestic and export product categories would render the results questionable.

12. In principle, problems of multicollinearity may emerge between the education and GDP per capita variables. All the same, given the broader meaning of the GDP variable, we keep both variables in the equation.

13. We also included a dummy variable for the existence of nontariff import barriers on used machinery, but we dropped it from the equation because the sign of the estimated coefficient was inconsistent and the coefficient was never significant. The lack of significance may have been due to incomplete data. The variable was generated using reports of the existence of nontariff barriers on used machinery in various surveys of trade policy in the countries in the sample, but some barriers may not have been reported.

progress embodied in new machinery and skill constraints faced by firms in developing countries. We tested hypotheses based on the model using data on U.S. exports of new and used metalworking machinery, differentiated by country of destination.

The results tend to corroborate the view that demand for used equipment is relatively high in lower-income developing countries. The proportion of each type of machinery bought secondhand is especially high for high-tech equipment requiring more sophisticated operating skills. Econometric tests of the determinants of the trade in used machinery indicate a significant role for relative prices, but also for technological and skill factors. Our results provide some support for the hypothesis that the absorptive capacity of a country (the ability to master a new technology) affects the choice of type and vintage of machines.

The finding that developing countries buy a larger share of old vintage machines when machines have a high rate of technological progress may be of some concern because it implies that the technological gap between the North and the South is likely to increase with faster technological progress. What policies can reduce the risk of a growing technology gap? The traditional criticism of restrictions on imports of used equipment is that such restrictions deny firms access to more appropriate labor-intensive techniques embodied in older vintage machinery. Our model and results indicate that, in addition, such restrictions force firms to buy more expensive new equipment that they may not be able to operate at full efficiency because of skill constraints in the labor force. Instead of imposing restrictions on used machinery imports, countries should concentrate on fostering education (particularly technical education) in an effort to improve the overall investment environment and firms' capacity to absorb new technologies.

APPENDIX: DATA DESCRIPTIONS AND SOURCES

Variable	*Description*	*Source*
Q_{ij}	Quantity of used machinery of type *i* exported to country *j* as a proportion of total machinery of type *i* imported by country *j*, 1990–94	U.S. Department of Commerce, Bureau of the Census, Exports of Merchandise CD-ROM
w_j	Average annual wage in country *j* in dollars, 1990–94	United Nations Industrial Development Organization, *Industrial Statistics*
r_j	Average real interest rate in country *j*	International Monetary Fund, *International Financial Statistics*
T_{ij}	Depreciation of machinery of type *i* imported by country *j*, measured as the difference between the logs of unit values of new and used machinery	Calculated from export data from U.S. Department of Commerce, Bureau of the Census, Exports of Merchandise CD-ROM
S_{ij}	Skill requirement index for machines of type *i* exported to country *j*	Based on classification by Anthony Bratkovich, engineering director of the U.S. Association for Manufacturing Technology
Y_{ij}	GDP per capita in dollars in country *j* (1990–96 average)	World Bank data
E_j	Education level in country *j*, measured by average years of school (over 1990–96)	World Bank database (STARS)
t_{ij}	Tariff rate on imports of machinery of type *i* imposed by country *j*	World Bank data

REFERENCES

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Bardhan, Pranab. 1970. *Economic Growth, Development, and Foreign Trade: A Study in Pure Theory.* New York: John Wiley and Sons.

Benhabib, Jesse, and Aldo Rustichini. 1991. "Vintage Capital Investment and Growth." *Journal of Economic Theory* 55(December):323–39.

Bond, Eric. 1983. "Trade in Used Equipment with Heterogeneous Firms." *Journal of Political Economy* 91(4):688–705.

Chari, V. V., and Hugo Hopenhayn. 1991. "Vintage Human Capital, Growth, and the Diffusion of New Technology." *Journal of Political Economy* 99(6):1142–65.

Dasgupta, Partha, and Joseph Stiglitz. 1988. "Learning by Doing: Market Structure and Industrial and Trade Policies." *Oxford Economic Papers* 40(June):246–68.

David, Paul A. 1993. "Knowledge, Property, and the System Dynamics of Technological Change." In *Proceedings of the World Bank Annual Conference on Development Economics 1992.* Washington, D.C.: World Bank.

Evenson, Robert, and Larry Westphal. 1994. "Technological Change and Technology Strategy." In T. N. Srinivasan and Jere Behrman, eds., *Handbook of Development Economics*. Vol. 3. Amsterdam: North-Holland.

Gabisch, Gunter. 1975. "A Vintage Capital Model of International Trade: The Case of Trade with Second Hand Machines." *Journal of International Economics* 5(November):365–83.

James, Dilmus. 1974. *Used Machinery and Economic Development, 1984*. MSU International Business and Economic Studies. East Lansing: Michigan State University.

Jovanovic, Boyan, and Glenn MacDonald. 1993. "Competitive Diffusion." NBER Working Paper 4463. National Bureau of Economic Research, Cambridge, Mass. Processed.

Jovanovic, Boyan, and Yaw Nyarko. 1995. "The Transfer of Human Capital." *Journal of Economic Dynamics and Control* 19(June):1033–64.

———. 1998. "How Does Productivity Depend on Past R&D?" In Giorgio Barba Navaretti, Partha Dasguptu, Karl G. Müler, and Domenico Siniscalco, eds., *Creation and Transfer of Knowledge: Institutions and Incentives*. Heidelberg and Berlin: Springer Verlag.

Note di Lavoro. Fondazione Eni Enrico Mattei, Milan. Processed.

Keller, Wolfgang. 1994. "Absorptive Capacity: Understanding the Creation and Acquisition of Technology in Development." Department of Economics, Yale University, New Haven, Conn. Processed.

Lall, Sanjaya. 1987. *Learning to Industrialise: The Acquisition of Technological Capabilities in India*. London: Macmillan.

Mainwaring, Lynn. 1986. "International Trade in New and Used Machines." *Cambridge Journal of Economics* 10(September):247–63.

Oliner, Stephen D. 1993. "New Evidence on the Retirement and Depreciation of Machine Tools." Working Paper Series 147. Board of the Federal Reserve System. December. Processed.

Pack, Howard. 1977. "The Optimality of Used Equipment: Calculation for the Cotton-Textile Industry." *Economic Development and Cultural Change* 26(2):307–25.

Pack, Howard, and Michael Todaro. 1970. "Industrialization, Employment, and the Choice of Alternative Vintage Equipment in Less Developed Countries." Center Discussion Paper 5. Economic Growth Center, Yale University, New Haven, Conn. Processed.

Schwartz, Sandra. 1973. "Second Hand Machinery in Development or How to Recognize a Bargain." *Journal of Development Studies* 9(4):544–55.

Sen, Amartya K. 1962. "On the Usefulness of Used Machines." *Review of Economics and Statistics* 44(3):346–48.

Smith, M. A. M. 1974. "International Trade in Second Hand Machines." *Journal of Development Economics* 1(3):261–78.

———. 1976. "International Trade Theory in Vintage Models." *Review of Economic Studies* 43(January):99–113.

Thoumi, Francisco. 1975. "Theory of the Optimum Age to Import a Durable Good." *Journal of Development Economics* 2(June):145–60.

Wonnacott, Paul. 1994. "The Automotive Industry in Southeast Asia: Can Protection Be Made Less Costly?" Institute for International Economics, Washington, D.C. September. Processed.

THE WORLD BANK ECONOMIC REVIEW, VOL. 14, NO. 1: 111-27

An Argument for Deregulating the Transfer of Agricultural Technologies to Developing Countries

David Gisselquist and Jean-Marie Grether

In the past few decades many developing countries have liberalized trade and investment, removing barriers to imports and allowing the introduction of new foreign technologies. Unfortunately, agriculture often remains outside this reform process. Regulatory obstacles continue to restrain the transfer of technologies through private trade in seeds and other inputs. Industrial countries characteristically allow the transfer of private and public technologies through multiple channels. Developing countries often force technology transfer through a single channel controlled by government agencies, with an emphasis on official performance tests. This article analyzes the institutional arrangements governing the international transfer of new agricultural technologies, examining the cases of agricultural machinery in Bangladesh and seed varieties in Turkey. The analysis shows that allowing the private transfer of technologies and refocusing input regulations on externalities could lead to significant productivity and income gains in developing countries.

In the past several decades many developing countries have liberalized their trade and investment regimes by eliminating barriers to imports and allowing in new foreign technologies. Some recent models suggest that the efficiency gains from relaxing trade restrictions on production inputs could be substantial (for example, Romer 1994). Unfortunately, agriculture often remains outside this reform process. Regulatory obstacles continue to impede the transfer of technologies through private trade in seeds and other inputs. With limited access to new technologies, many farmers in developing countries continue to rely on traditional or old crop varieties, inefficient livestock breeds and feeding technologies, and older and more dangerous pesticides. Barriers to the introduction of agricultural technologies are particularly worrisome for low-income countries that depend heavily on agriculture. And removing existing barriers may lead to substantial productivity and income gains. For example, in industrial countries plant breeding boosts poten-

David Gisselquist is an independent consultant, and Jean-Marie Grether is a lecturer at the University of Geneva and the University of Neuchâtel. Their e-mail addresses are *dgisselquist@worldbank.org* and *grether@ibm.unige.ch@internet*. The authors thank Jaime de Melo, three anonymous referees, and participants at the conference Trade and Technology Diffusion: The Evidence with Implications for Developing Countries, sponsored by the Fondazione Eni Enrico Mattei and the International Trade Division of the World Bank and held in Milan on April 18–19, 1997, for helpful comments and suggestions.

tial maize yields about 0.7 percent a year; in developing countries, where breeding has been less intense to date, the potential impact of new maize hybrids can be much greater.

This article addresses two questions regarding the transfer of technologies in developing countries: what are the institutional obstacles to the transfer of technologies mediated through the market, and are they based on convincing arguments? Apart from general import and investment barriers, whose consequences have been widely documented, many governments impede the private transfer of technologies in agriculture with specific regulations that block the sale of inputs—whether imported or domestically produced—except for technologies that they have approved after a review process that includes official performance tests. The impact of performance-based input regulations on productivity and income has not been measured for developing countries. If such regulations were benign or beneficial, countries that deregulated would not show any changes in technology flows, input sales, or productivity growth.

This article looks at two case studies—agricultural machinery in Bangladesh and seed varieties in Turkey. It asks whether regulatory reforms brought dramatic gains in technology transfer, input trade, and productivity. We note that markets should not be totally liberalized because any policy reform must explicitly consider safeguards to limit externalities. Specifically, we do not challenge performance-based tests for conventional pesticides and other inputs with substantial public health or environmental externalities.

I. The Generation and Transfer of Technologies in Agriculture

Agriculture has become a high-technology field, with rapid advances in crop and livestock genetics, pest and livestock management, and machinery. For many field crops the average market life of a variety is no more than five to seven years, and for vegetables it can be as short as two years. The use of conventional pesticides—broad-spectrum poisons—is giving way to a range of relatively low-risk pest management techniques (for example, insect growth regulators, pheromones, microbial pesticides, and inoculants).

Concurrently, private research has become more important. In the United States the private share of agricultural research expenditures, which was less than 50 percent in the mid-1980s, increased to 56 percent in 1992 (Clive 1996). Estimates for 1993 show that in most industrial countries the private sector provided more than 50 percent of agricultural research funds. This share was highest in the United Kingdom, at 62 percent (Alston, Pardey, and Roseboom 1998). In some areas, such as biotechnology, most research is financed by private sources.

Public research in industrial countries has responded to increases in private research by moving upstream into basic research that private companies can build on to develop new technologies for market application. Also, universities and other public research organizations have been applying for intellectual property rights and then selling or licensing new technologies to private companies. The

global trend toward stronger intellectual property rights, as evident in the World Trade Organization agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS), promises further increases in the private sector's share of agricultural research and the generation and transfer of technologies.

The recent upsurge in private sector research comes at a time when many donors and developing-country governments have been restraining—if not cutting—their budgets for international and national public agricultural research. Threats to public research in developing countries raise strong concerns among agricultural experts. Alston, Pardey, and Roseboom (1998) note that private research spending in developing countries is seldom more than 10–15 percent of total agricultural research budgets. Returns to some types of agricultural research—for example, crop management to reduce losses to pests—are not appropriable, and hence private companies underinvest in them. Furthermore, evidence accumulated over decades shows high social returns to public investment in agricultural research (see Alston and others 1998).

The debate about the appropriate role of the state in agricultural research is background to this article, particularly the sections discussing the institutional framework (section II) and policy implications (section IV). For example, Byerlee (1998) discusses a new paradigm for national agricultural research systems, and Pray and Umali-Deininger (1998) assess the ability of the private sector to compensate for declines in public funding. However, the scope of our study is limited to international technology transfer, with a particular emphasis on performance tests. Taking the above trends as given, we focus on the specific problem of farmers' access to new technologies, whether they originate from public or private research.

The issue of access to foreign technologies is particularly important because, as in other high-technology fields, agricultural technology is now international. Leading countries continually borrow and build on research from other countries. Technologies move among people, organizations, and countries through publications, discussions, licensing agreements, and international sales. Despite some instances of institutional collaboration (for example, between international centers within the Consultative Group on International Agricultural Research, CGIAR), there is no master plan for research and dissemination. Coordination is achieved mainly through communication and marketing.

The mode of international technology transfer differs according to the types of inputs. Direct transfers through imports are common for high-value seeds, such as hybrid vegetables, new proprietary pesticides, and engine-powered machinery. In-country production using an imported design developed from foreign research also is common, particularly in large countries, for seeds of field crops (for example, a company may import breeders' seed to multiply in-country), out-of-patent pesticides, vaccines, or simple livestock feed additives. Finally, local research may adapt foreign technologies (for example, germ plasm or machinery) to local conditions; this can be particularly important for some categories of technology that depend on local crop management practices, such as integrated pest management.

Whatever the source, most new technologies reach farmers through marketed inputs. New varieties are embodied in seeds, new pest management technologies in pesticides or spraying equipment, and new feed technologies in pre-mixes. Companies extend technologies to farmers through test plots, demonstrations, and dealers. However, the diffusion of technologies has been uneven. Many developing countries lag behind, in part because of self-imposed barriers to the introduction of private agricultural technologies.

II. Channels for Transferring Technologies

Although regulations governing the transfer of agricultural technologies and the trade of inputs vary widely across countries, it is useful to identify two stylized patterns. Access to new technologies may come either through multiple channels, with regulations focusing on externalities, or through a single channel, with regulations based on performance tests. After characterizing this basic distinction, we discuss why the single-channel system, in spite of its restrictiveness, is still in force in many developing countries.

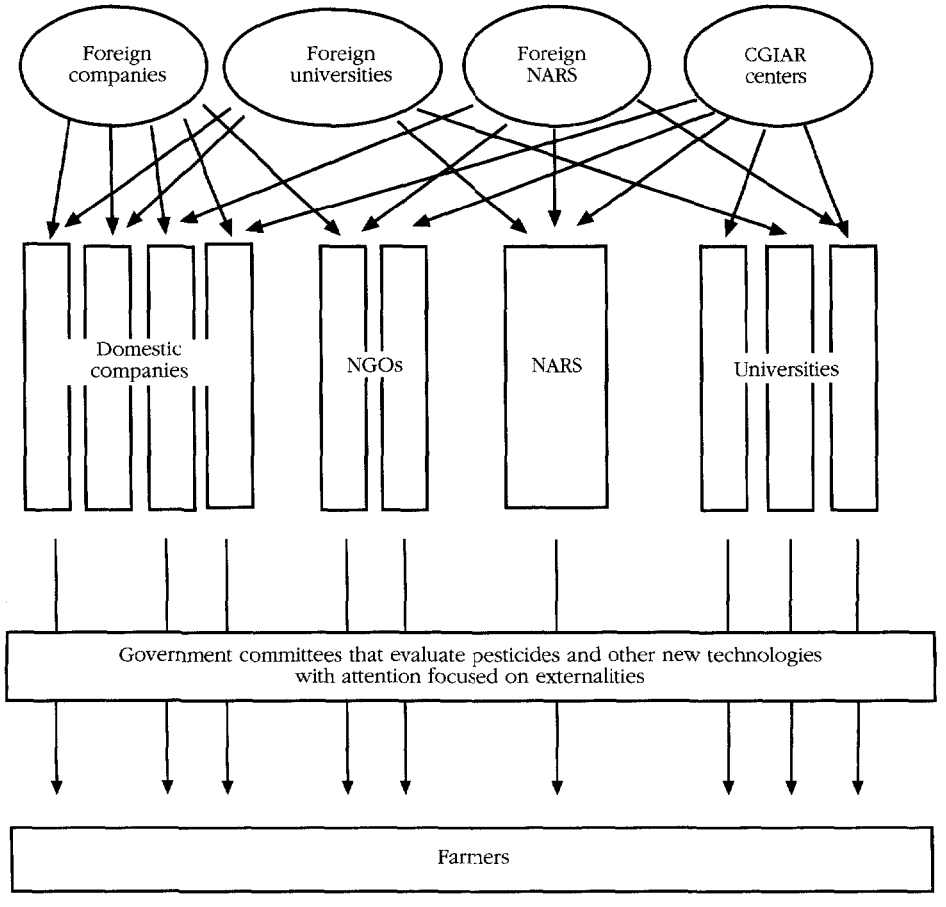
Two Institutional Frameworks

In industrial countries (and in some developing countries) governments maintain liberal trade regimes for foreign and domestic inputs, allowing multiple channels for the introduction of new technologies. Companies, universities, nongovernmental organizations, and government research institutes introduce new inputs that embody new technologies from their own or foreign research (figure 1). Governments regulate inputs to limit externalities (for example, by not approving dangerous pesticides), but otherwise allow companies to market new technologies, trusting that farmers and companies interacting through markets will be able to chose those that are most efficient.

Although all industrial countries allow multiple channels for introducing technologies, there are some differences in regulatory systems. In the European Union, for example, each member government tests plant varieties for performance and automatically accepts varieties approved by any other European Union government without further tests. By contrast, the United States, India, and some other countries allow the sale of seed without variety registration or official performance tests; registration is available, but optional. Despite some differences and special cases, regulatory systems in industrial countries share the same underlying logic, allowing markets to evaluate performance, while focusing regulations on externalities. This liberal approach to technology transfer is appropriate for agriculture, a field with rapid technical change and for which local conditions are critical in shaping the impact of new technologies.

In contrast to the market-friendly regulatory regimes common in industrial countries, many developing and transition countries strictly limit market access for new agricultural technologies. Restrictions are most common and problematic for seeds, but they also may interfere with machinery, fertilizers, low-risk

Figure 1. *Multiple Channels for New Agricultural Technologies*



Note: NARS are national agricultural research systems. CGIAR is the Consultative Group on International Agricultural Research. CGIAR centers are the International Rice Research Institute (IRRI) and other international agricultural research centers associated with the CGIAR. NGOs are nongovernmental organizations.

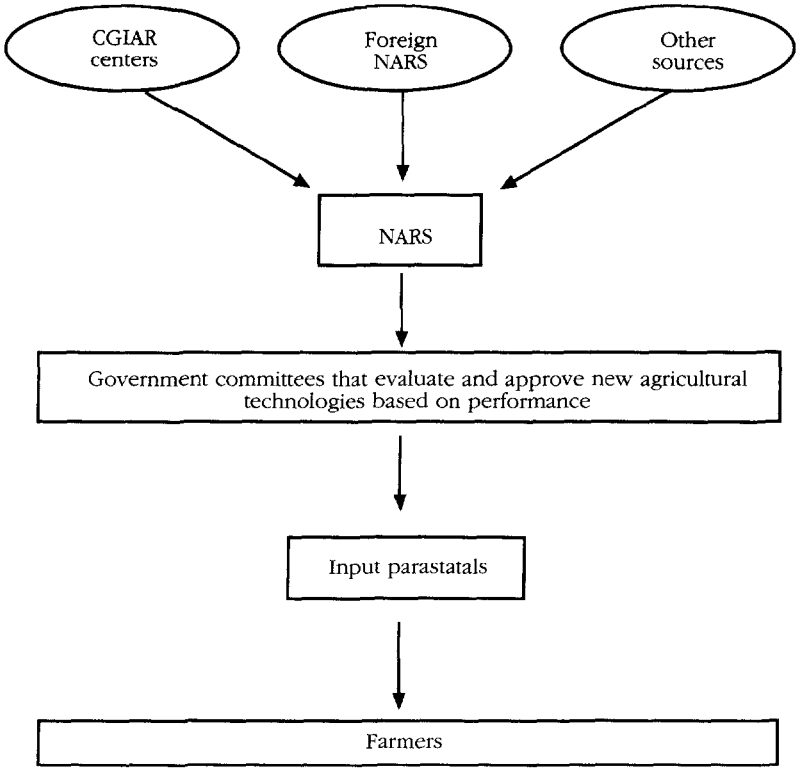
pesticides, feed mixes, and other items. Many developing countries maintain positive lists of allowed inputs, even those for which externalities are not a serious concern. For example, many countries list allowed plant varieties, and some also list allowed models of machinery, compositions of fertilizers, and feed mixes based on official performance tests. Bhagwati (1978) and Krueger (1978) describe positive lists, which are common in foreign trade regimes that use quantitative restrictions. Positive lists are far more restrictive than negative lists, which allow anything not listed.

For most inputs environmental or public health concerns can be addressed with negative lists. For example, instead of approving each feed mix, govern-

ments can simply ban or limit potentially dangerous components, such as hormones. Although positive lists are standard across all countries for pesticides and veterinary medicines, countries differ in the conditions for registering new products. Many developing countries limit new products in ways that restrict competition and entry. For example, in 1984 Indonesia decreed that no more than three companies could sell pesticides with any particular active ingredient.

With regulations and policies that make it difficult for companies to enter and to operate, many developing countries effectively block almost all transfers of private technologies for seeds and other major categories of agricultural inputs. Thus government agencies are left as the dominant or only channel for technology transfer (see figure 2). Under this system a centralized government research establishment identifies new varieties and other new technologies, parastatals produce or import and sell inputs, and extension agents encourage farmers to take what is offered. Often a single-channel system operates for some, but not all, inputs.

Figure 2. *Single Channel for New Agricultural Technologies*



Note: NARS are national agricultural research systems. CGIAR is the Consultative Group on International Agricultural Research. CGIAR centers are the International Rice Research Institute (IRRI) and other international agricultural research centers associated with the CGIAR.

A single-channel system severely constrains the flow of new technologies. In many developing countries with single-channel systems farmers are offered an average of less than one new seed variety a year for each major crop, while farmers in countries with multiple channels may see dozens of new varieties each year for a single major or minor crop or vegetable. Even where private companies are able to operate, regulatory costs limit the transfer of private technologies. Regulatory costs are particularly troublesome in small markets—small countries or minor crops—where companies may judge that registering a new technology is not worth the effort, leaving farmers with no access.

Arguments for Single-Channel Systems

Many developing-country governments and donor experts continue to promote, endorse, or accept single-channel systems and performance-based regulation of inputs, despite their limitations. The standard argument sustaining this position is that deregulation would lead to the diffusion of unsatisfactory inputs because farmers either lack the necessary information to discriminate or are susceptible to false marketing. Thus farmers should be protected from those who "might try to market an unsatisfactory variety simply to recoup breeding costs" (Kelly 1989: 43). Apart from underestimating the capacity of farmers to learn about inputs at a low cost, this argument is debatable on several counts. If lack of information is the issue, there certainly are more efficient ways to address it than through a registration process that limits the range of available inputs. Moreover, fears that giving farmers more choice will encourage the spread of inferior varieties—or fertilizer compositions or feed mixes—are not supported by experiences in India and other countries that allow the sale of untested seed varieties.

In some cases, such as that of low-risk pesticides, excessive reliance on performance tests may even exacerbate externalities. Conventional broad-spectrum poisons, which kill a wide range of pests outright, but which also threaten public health and cause environmental damage, are easy to manage in efficacy tests and have large markets for multiple pests. Nontoxic biopesticides that interfere with insect mating or maturation can be more difficult to manage, so that results from efficacy tests are misleading. In addition, they may be limited to specific insects and crops. Forcing low-risk pesticides to pass expensive efficacy tests can leave farmers with limited access to low-risk products (see Benbrook and Gisselquist 1996).

Many donor experts see deregulation to allow the introduction of private technologies as either irrelevant or threatening. Evenson and Westphal (1995: 2248–49) assert, "Important forms of adaptive agricultural R&D require a substantial commitment of resources dedicated to developing techniques for a particular set of agronomic conditions." This line of reasoning leads to the conclusion that deregulation has minimal impact because the private sector does little research in developing countries and offers few suitable technologies. However, casual observation shows that pesticides, fertilizers, machinery, livestock feeds, and breeds can move across borders. Recent studies of multinational adoption of CGIAR rice

and wheat lines bred in Mexico and the Philippines show that crop varieties also move across borders (Maredia, Ward, and Byerlee 1997; Evenson and Gollin 1997).

Another common reservation about deregulation among donor experts is that private technology transfer might somehow undermine public funding or public research. This reservation is difficult to argue against because it is not rational. A decision to allow more private technologies through deregulation is essentially costless to the government. It would be entirely independent of any other government or donor decision about the extent of funding for public research. Even if deregulation leads to private technologies taking over some agricultural activity (for example, private fertilizer dealers could design compounds for particular crops and regions), public scientists could continue to work on whatever projects the government and donors support (for example, testing soils and fertilizers, presenting farmers with competing advice). If industrial countries serve as a guide, systems with multiple channels have more advanced agricultural technologies and higher levels of public research as a share of agricultural gross domestic product (GDP) than single-channel systems. Hence, to think of public research and private technology as an either-or choice is misleading; they are better thought of as complements.

Persistent Regulation of Technology Trade

The case for single-channel systems seems to be weak. However, single-channel systems have significant support among developing-country governments. This persistence may be explained by the influence of the beneficiaries of such systems. In some cases private monopolists or oligopolists may oppose opening the door to competition. An example is the Bangladeshi company that was the exclusive distributor of the only Chinese power tiller approved for import. The company suffered in 1989 when reforms allowed competing low-cost models to enter the market (see section III). Similarly, pesticide companies in Zimbabwe benefit from regulations requiring three years of (redundant) in-country efficacy tests for competing brands of old and out-of-patent pesticides. And in several Central European countries spokespersons for seed companies that have large market shares privately favor the compulsory registration of varieties to hold off competing firms.

However, the strongest call for governments to limit the transfer of private technologies characteristically comes from public sector scientists. Public sector scientists can be influential far beyond their numbers. They are the gatekeepers in a single-channel system. Often prestige and concerns about job security seem to be the major issues, although some scientists may also receive monetary benefits from testing fees or from bribes.

In countries with single-channel systems public sector scientists not only have blocked private technologies, but also have blocked technologies from CGIAR centers. For example, for several years in the 1960s government scientists in India and Turkey refused to approve dwarf wheat varieties from the International Maize

and Wheat Improvement Center (CIMMYT), despite clear evidence of superior yields from field tests. In India government scientists reportedly falsified field test results to avoid having to approve CIMMYT wheat. Presumably, the scientists wanted to breed their own varieties, crossing CIMMYT and local wheat. However, breeding takes time; regardless of what is planned for local breeding, it makes sense to give farmers access to good varieties that are already available. In both countries donors helped to bring the issue to the attention of senior government officials, who pushed through the approval of CIMMYT varieties against domestic objections.

Bowing to the wishes of scientists in national agricultural research systems, the International Rice Research Institute (IRRI) and other CGIAR centers have adopted and maintain the fiction that they do not release varieties, but rather lines, which the national agricultural research systems then further develop into varieties. It is true that research systems play a large and positive role in introducing and adapting CGIAR technologies. However, many CGIAR lines are suitable for release and are widely released by the research systems after adaptive trials only, without further breeding. Incredibly, most CGIAR lines are illegal in many if not most developing countries—because scientists from the national agricultural research systems have not approved them for lists of registered and allowed varieties.

Donors also have supported single-channel systems, but their logic and motives are not clear. Allowing the introduction of private technologies in agriculture is consistent with widely accepted patterns in other sectors, such as computers and software, where private technologies move without government approval. However, donors have a history of paying special attention to agriculture, including a long involvement with public sector scientists in CGIAR centers and national agricultural research systems. Recently, public agricultural research has been under attack from budget cutters, despite continuing high returns to investment. Maintaining existing public research capacity is thus a challenge that legitimates and absorbs a significant share of donor effort and attention. But here again, there should be no confusion between the necessity of preserving public research and the gains from deregulation.

III. The Impact of Deregulation: Two Case Studies

Although the effects of trade liberalization are widely documented, there is little evidence on the impact of performance-based regulations. Two studies of industrial countries suggest that losses in terms of forgone income may be huge for cotton in California (Constantine, Alston, and Smith 1994) and for wheat in Canada (Ulrich, Furtan, and Schmitz 1987). To our knowledge, no similar estimates exist for developing countries. However, based on a sample of 50 mostly developing countries, Pray and Echeverría (1988) find that seed imports and private research are significantly correlated with maize yields.

One research strategy to measure the impact of deregulating inputs is to work with aggregate data, comparing levels or growth rates of agricultural production

or total factor productivity across countries or over time. The most convincing cross-country argument is trivial: no research is required to establish that industrial countries as a group have more open and competitive input industries and higher levels of agricultural technology and production than low-income developing countries. However, it is difficult to interpret cross-country comparisons of growth rates because countries with restrictive systems can show fast growth from a low and constrained base (resulting in part from many years of obstructed technology transfer). Also, regulatory systems are varied and complex, so it takes time to understand what is going on in any one country; the result may be difficult to express numerically for econometric analysis. Another set of problems undermines intertemporal models that use aggregate data: changes in the regulation of inputs often occur in conjunction with other macroeconomic and microeconomic reforms. It can be difficult to disentangle the impacts of these reforms, along with climate and other factors, from the impact of regulatory reforms.

In this section we trace the impact of regulatory changes in Bangladesh and Turkey on technology transfer for a particular input, then on input sales, and finally on specific agricultural activities or outputs. These case studies are not designed to measure the aggregate impact of regulatory reforms on all of agriculture, but rather to show that reforms have a positive and not insignificant impact on selected outputs or activities. Therefore, if production increases with deregulation, then regulations on balance do not protect farmers from inferior technologies, and regulations are not irrelevant.

Agricultural Machinery in Bangladesh

The case of agricultural machinery in Bangladesh before reform constitutes an almost ideal illustration of the single-channel system. Prior to reform in the late 1980s, the Ministry of Agriculture listed "standardized" (tested and approved) models of diesel engines for irrigation and power tillers. The Bangladesh Agricultural Development Corporation, a parastatal, imported listed engines and sold them at subsidized prices. Models not on the list could not be imported.

Reform arrived in 1988–89. Along with modest tariff cuts on diesel engines for irrigation (from 15 to 0 percent), the Ministry of Agriculture did away with lists of engines standardized for irrigation, allowing all models. Subsidized parastatal sales continued for several years, but farmers increasingly shifted to private traders, who offered convenience along with a wide range of low-cost models. By the end of 1991 private traders clearly dominated the market. Regulatory reform—allowing the private import and sale of new and less-costly models of diesel engines and power tillers from China—was followed by a sharp increase in the sale and use of imported machinery.

Regulatory reform allowed farmers to choose cheaper equipment (see table 1). The retail cost for the most common minor irrigation investment (a 12 horsepower engine and a 100 millimeter diameter tubewell for lifting groundwater) fell below $500 by the end of the 1980s, less than half what it had been with

Table 1. *Impact of Deregulation of Imports of Agricultural Machinery in Bangladesh, 1988–96*

Type of input	*Fall in retail price*	*Impact on farm use*
Diesel engines for minor irrigation	More than 50 percent	Increase of 170 percent over eight years in the number of small pumps operating (extending new irrigation to an estimated 16 percent of gross cropped area)
Power tillers	More than 40 percent	Machinery cultivation extended from 0 percent in 1988 to 15–40 percent of cultivated area, depending on the season

Note: Regulatory reforms (1988–89) included trade liberalization of agricultural machinery and the suppression of compulsory registration.
Source: Government of Bangladesh, Ministry of Agriculture (1995a, b) and authors' calculations.

subsidies in 1981–82. After the reform minor irrigation expanded at record rates. From 1988 to 1996 the number of small power pumps lifting ground or surface water for irrigation increased by 170 percent, or 390,000 units, delivering new irrigation to roughly 16 percent of the gross cropped area (assuming that each new pump irrigated an average of 4 hectares). Markets also moved toward smaller equipment (4–8 horsepower engines and 75–100 millimeter diameter tubewells).

The 1988 pre-reform list of standardized power tillers included only one low-cost model from China (cost, insurance, and freight [CIF] import price about $1,000), one from the Republic of Korea (CIF about $1,700), and about 10 others from high-cost sources (CIF well over $2,000). Dealers for the two low-cost models dominated trade, taking advantage of limited competition to sell them for more than $2,000; with no duties in late 1988, dealers realized large profits. When standardization ended, multiple models from China, with CIF near $1,000, entered the market, and competition soon cut the retail price to about $1,300. Before the reform power tillers were so rare that an observer typically would not see any during a multi-day tour of rural areas. By 1996 power tillers prepared an estimated 15–40 percent of the land for cultivation, depending on the season (based on power tiller imports and assuming that each power tiller lasts five years and cultivates 25 hectares in a season).

Seed Varieties in Turkey

A 1963 Turkish seed law gave the Ministry of Agriculture authority over seed production and trade. Through regulations based on the law, the Ministry made variety registration and in-country performance tests compulsory for most crops, set seed prices annually, and extended import and export controls well beyond phytosanitary concerns. In practice, the Ministry limited the approval of varieties for most field crops to those sponsored by government research agencies; for vegetables, it allowed only a limited range of private varieties.

At the beginning of the 1980s difficulties with Turkey's single-channel system included widespread smuggling of vegetable seeds, failure to popularize hybrid maize, and expensive government agencies serving no more than 10 percent of the planted area. The government consequently revised seed policies to encourage private participation in seed production and trade. Between 1982 and 1984 the government removed seed price controls, relaxed foreign investment controls, and eased (but did not entirely dismantle) compulsory registration of varieties by reducing testing requirements and allowing private companies to conduct their own tests.

Reforms brought large increases in the number of varieties allowed for sale, as well as rapid expansion of participation by private companies (see table 2). Between 1982 and 1994 the number of allowed varieties increased 670 percent for hybrid maize, 2,400 percent for hybrid sunflowers, and 3,400 percent for soybeans. Most of these new varieties were direct transfers, often from parts of Western Europe and the United States that shared the same latitudes. Most were proprietary varieties, although some came from foreign or international public research. As a result, the share of commercial seed sales through private companies soared—in 1993 exceeding 90 percent for maize and sunflower hybrids, soybeans, and potatoes. The number of private companies increased from about 5 to 80 in 1980–94. During this period most major seed multinationals established a presence in Turkey through subsidiaries, joint ventures, or licensing agreements.

Trade reforms led to an initial increase in seed imports for some crops. For example, imports of hybrid maize seed exceeded production in 1985. Soon, however, local seed production expanded to meet local demand and then pushed into export markets as well. Exports of hybrid maize seed exceeded imports in 1988 and reached a quarter of total production in 1992. Similar trade shifts occurred for hybrid sunflower seed. Once reforms allowed seed technologies to enter, Turkey was able to exploit its comparative advantages of a good climate, developed scientific skills, and low labor costs.

Available data on maize yields allow for a rough estimate of the gains from private hybrids following the reforms. Gisselquist and Pray (1997) estimate a yield response function over 1961–91 (see table 3). Regressors include the percentage of maize area sown in private hybrids, annual fertilizer use, annual rain-

Table 2. *Impact of Deregulation of Trade in Seed in Turkey, 1982–94*

Crop	*Harvested area, 1990 (hectares)*	*Varieties available, 1982*	*New varieties introduced, 1982–94*	*Private share of commercial seed production (percent)*	
				1985	*1994*
Wheat	9,400,000	21	62	0.5	8.8
Hybrid sunflowers	715,000	3	74	88.9	98.9
Cotton	841,000	9	19	0.0	0.1
Hybrid maize	155,000	24	185	85.7	97.3
Potatoes	192,000	31	51	11.3	91.7
Soybeans	74,000	2	70	42.1	94.7

Source: Government of Turkey, SIS (1992) and Government of Turkey (various years).

Table 3. *Maize Yield Response Function, Turkey, 1961–91*

Explained variable	*Share of hybrid planted area*	*Fertilizers per hectare*	*National rainfall*	*Trend*	*Adjusted* R^2
Maize yield (tons per hectare)	2.89	1.4E-03	4.58E-04	5.34E-02	0.924
	(0.52)	(5.3E-03)	(7.14E-04)	(2.73E-02)	

Note: Numbers in parentheses are standard deviations.
Source: Gisselquist and Pray (1997).

fall, and a trend variable to control for other factors (such as transport improvements and extension). Data on the irrigated area for maize were not available.

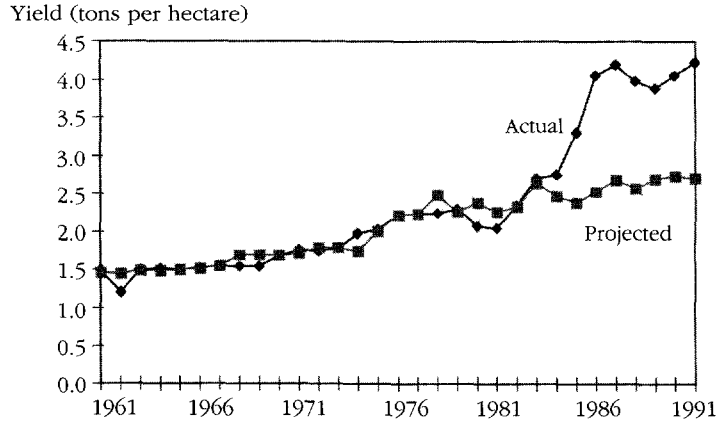
The explanatory power of the model is satisfactory. All coefficients exhibit the expected sign, but neither fertilizer nor rainfall is significant at the 95 percent level. Gisselquist and Pray simulate projected maize yields in the absence of reform using the estimated coefficients, but with zeros for post-reform hybrids. A comparison of actual yields and projected nonreform yields suggests that private maize hybrids boosted maize yields more than 50 percent (see figure 3).

Gisselquist and Pray use this result to estimate the magnitude of the income benefit from post-reform hybrid maize in Turkey. The increase in average net economic returns per hectare of maize during 1990–92 was $153, equivalent to 25 percent of the gross economic value (table 4). With a total maize area of 515,000 hectares, this result implies an annual net economic gain of $79 million for Turkey's maize farmers.

IV. Conclusions and Policy Implications

The evidence presented in this article suggests that deregulating the trade of inputs leads to significant increases in the range and quality of inputs available to

Figure 3. *Actual and Projected Maize Yields in Turkey, 1961–91*



Source: Gisselquist and Pray (1997).

Table 4. *Estimated Average Annual Net Economic Benefits from Hybrid Maize, Turkey, 1990–92*

Indicator	*Volume*[a]	*Change in volume due to hybrids*[a]	*Unit price*[b]	*Change in value due to hybrids (dollars per hectare)*
Average yield	4.13	1.43	148	211
Projected yield without hybrids	2.7[c]	n.a.	n.a.	n.a.
Cost				
Nonhybrid seed	37	−11.1[d]	0.148	−2
Hybrid seed	28	8.4[d]	2.97	25
Harvesting and drying[e]	n.a.	n.a.	n.a.	35
Net gain	n.a.	n.a.	n.a.	153

n.a. Not applicable.
a. Yields are in tons per hectare, and costs are in kilograms per hectare.
b. Yields are in dollars per ton, and costs are in dollars per kilogram.
c. Based on the yield response function described in the text.
d. Assumes that 30 percent of the maize area is planted in hybrids.
e. Estimated as one-sixth of the value of production.
Source: Gisselquist and Pray (1997).

farmers, which in turn raises productivity and income. From the case studies of Bangladesh and Turkey we recommend that countries with single-channel systems revise their regulations, especially performance-based regulations for approving new technologies, and move toward multiple channels for technology transfer. Evidence of large gains with a greater range of technical options agrees with results from trade models. Of course, our results would be more robust if we could support them with additional studies of similar reforms. Preliminary findings from ongoing World Bank studies of input reforms in Zimbabwe and India corroborate the conclusions and recommendations in this article.

The recommendation to reduce and revise regulations governing the trade of inputs should not be confused with full trade liberalization. Regulations to control negative externalities in terms of public health and environmental damage should be maintained and even reinforced in several countries. For example, in the case of new pesticides, governments might consider not only maintaining a positive list of allowed products, but also levying taxes on allowed pesticides, with rates determined according to externalities. More generally, existing regulations based on performance could be redesigned to focus on externalities (see table 5). For inputs with significant externalities, such as medium- and high-risk pesticides, performance could be taken into consideration in deciding whether the gains are worth the risks.

Serious consideration of regulatory reform requires time and effort to deal with specific situations and concerns. For example, even if people accept the principle of limiting regulations to externalities, in some situations they will debate the existence or magnitude of externalities.

Market power, another potential source of market failure, may be a concern in the reform process. The risk that monopolies and oligopolies will dominate

Table 5. *Reforming Regulation on the Transfer of Agricultural Technology*

Input	*Common regulatory barrier*	*Proposed reforms to focus regulation on externalities*
Seeds	Governments prohibit the sale of seed except for registered (approved) varieties based in part on performance tests	Allow the sale of seed without variety registration
	Governments block seed imports to protect domestic seed production (using unreasonable phytosanitary arguments or other nontariff barriers)	Focus phytosanitary controls on diseases that are present in the exporting country but not in the importing country and that threaten real economic damage
	Governments demand that companies submit samples of inbred lines before the governments will allow the sale of hybrid seed	Allow importation and sale without deposit of a seed sample; this regulation has nothing to do with externalities
Pesticides	Governments require in-country efficacy tests for new products before allowing them for sale	List allowed products; allow the sale of new no-risk or low-risk products without in-country efficacy tests
Fertilizers	Governments limit the types of fertilizers allowed for sale based on expert opinions about soil nutrient deficiencies	Allow companies to sell fertilizers with any combination of nutrients; enforce truth-in-labeling and ban dangerous impurities
Agricultural machinery	Governments limit imports to lists of approved models, basing approvals on performance tests	Allow importation of any model, leaving farmers to assess performance against cost and other factors
Livestock feed	Governments set minimum standards for various nutrients or components or require prior registration and approval for all feed mixes based on expert opinions	Allow companies to sell any combination of feed components without registration; enforce truth-in-labeling and ban or regulate feed additives with negative externalities (hormones and antibiotics)
Veterinary medicines	Governments do not allow the import of (private) vaccines, arguing that government production is adequate, diseases do not exist, or quality is not secure	List allowed products; allow private and competing products but regulate them to ensure quality (impurities in vaccines can spread other diseases)

input markets is more serious for small and low-income countries. For example, a minimum of 20–30 seed companies in a country may be required to ensure that farmers have access to world technologies for all crops through competitive seed markets. (Five or six seed companies ensure a competitive market for a single crop, but companies specialize, so several times that number is required to cover

all crops.) National seed markets in most Sub-Saharan African countries are not large enough to support competitive, modern seed markets. But regulatory reforms can allow national markets to merge into regional markets that are large enough to ensure a competitive supply of seed for minor and major crops. Policies favoring regional seed markets in Africa include voluntary registration of varieties and limits on seed import controls to realistic phytosanitary concerns, allowing varieties and seeds to move more easily across borders.

Deregulation should not jeopardize efforts to maintain and increase local public research and mastery of technologies. Public research has been and can be a source of useful technologies, often adapting and improving foreign technologies for local conditions. Furthermore, some classes of agricultural technologies are public goods (for example, integrated pest management techniques that do not use broad-spectrum poisons), for which the social returns to research far exceed what can be recouped through the sale of inputs. Also, local research capacity strengthens the bargaining power of the country if prices for technology transfer are not competitive (Pack and Westphal 1986) and avoids a widening of the technological gap when spillovers from research and development are national rather than international (Grossman and Helpman 1991).

There may be good reasons to increase public research as the transfer of private technologies expands. The entry of more research-intensive input companies creates a competitive market for scientists and ideas, which improves the environment for managing public research. Research-intensive companies often work closely with the public sector and request more public research. And a commitment to boost funds for public research may help to blunt opposition to reform from government scientists who lose monopoly control over the transfer of technologies.

REFERENCES

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Alston, J. M., M. C. Marra, Philip Pardey, and T. J. Wyatt. 1998. *Research Returns Redux: A Meta-Analysis of Agricultural R&D Evaluations.* Environment and Production Technology Division Discussion Paper 38. Washington D.C.: International Food Policy Research Institute.

Alston, J. M., Philip Pardey, and Johannes Roseboom. 1998. "Financing Agricultural Research: International Investment Patterns and Policy Perspectives." *World Development* 26(6):1057–71.

Benbrook, Charles, and David Gisselquist. 1996. *Technology Transfer, Competition, and Trade Liberalization for Low-Risk Pest Management.* UNDP–World Bank Trade Expansion Program Occasional Paper 11. Washington D.C.: United Nations Development Programme and World Bank.

Bhagwati, Jagdish. 1978. *Foreign Trade Regimes and Economic Development: Anatomy and Consequences of Exchange Control Regimes.* Boston: Ballinger.

Byerlee, Derek. 1998. "The Search for a New Paradigm for the Development of National Agricultural Research Systems." *World Development* 26(6):1049–55.

Clive, James. 1996. "Agricultural Research and Development: The Need for Public-Private Partnership." Issues in Agriculture 9. Consultative Group for International Agricultural Research, Washington D.C. Processed.

Constantine, J. H., J. M. Alston, and V. H. Smith. 1994. "Economic Impacts of the California One-Variety Cotton Law." *Journal of Political Economy* 102(10):951–74.

Evenson, R. E., and Doug Gollin. 1997. "Genetic Resources, International Organizations, and Rice Variety Improvement." In David Gisselquist and Jitendra Srivastava, eds., *Easing Barriers to Movement of Plant Varieties for Agricultural Development.* World Bank Discussion Paper 367. Washington, D.C.: World Bank.

Evenson, R. E., and Larry Westphal. 1995. "Technological Change and Technological Strategy." In Jere Behrman and T. N. Srinivasan, eds., *Handbook of Development Economics.* Vol. 3A. Amsterdam: North-Holland.

Gisselquist, David, and Carl Pray. 1997. "The Impact of Turkey's 1980s Seed Regulatory Reform." In David Gisselquist and Jitendra Srivastava, eds., *Easing Barriers to Movement of Plant Varieties for Agricultural Development.* World Bank Discussion Paper 367. Washington, D.C.: World Bank.

Government of Bangladesh, Ministry of Agriculture. 1995a. *National Minor Irrigation Census, Bangladesh: 1994/95 Irrigation Season Main Report.* Dhaka.

———. 1995b. "Weekly Report on National Minor Irrigation Census, 1995/96, on 24/8/96." Dhaka. Processed.

Government of Turkey. Various years (1982–94). *Resmi Gazete.* Various issues.

Government of Turkey, SIS (State Institute of Statistics). 1992. *1990 Agricultural Structure and Production.* Ankara.

Grossman, Gene, and Elhanan Helpman. 1991. *Innovation and Growth in the Global Economy.* Cambridge, Mass.: MIT Press.

Kelly, A. Fenwick. 1989. *Seed Planning and Policy for Agricultural Production: The Roles of Government and Private Enterprise in Supply and Production.* London: Bellhaven.

Krueger, Anne. 1978. *Foreign Trade Regimes and Economic Development: Liberalization Attempts and Consequences.* Boston: Ballinger.

Maredia, Mywish, Richard Ward, and Derek Byerlee. 1997. "Assessment of the International Transfer of Wheat Varieties." In David Gisselquist and Jitendra Srivastava, eds., *Easing Barriers to Movement of Plant Varieties for Agricultural Development.* World Bank Discussion Paper 367. Washington, D.C.: World Bank.

Pack, Howard, and Larry Westphal. 1986. "Industrial Strategy and Technological Change: Theory versus Reality." *Journal of Development Economics* 22(6):87–128.

Pray, Carl, and Rubén Echeverría. 1988. "Transferring Hybrid Maize Technology: The Role of the Private Sector." *Food Policy* 11(November):372–400.

Pray, Carl, and Dima Umali-Deininger. 1998. "The Private Sector in Agricultural Systems: Will It Fill the Gap?" *World Development* 26(6):1127–48.

Romer, Paul. 1994. "New Goods, Old Theory, and the Welfare Costs of Trade Restriction." *Journal of Development Economics* 43(2):5–38.

Ulrich, Alvin, William H. Furtan, and Andrew Schmitz. 1987. "The Cost of a Licensing System Regulation: An Example from Canadian Prairie Agriculture." *Journal of Political Economy* 95(1):160–78.

THE WORLD BANK ECONOMIC REVIEW, VOL. 14, NO. 1: 129-45

Geographical Targeting for Poverty Alleviation: An Introduction to the Special Issue

David Bigman and Hippolyte Fofack

In the face of rising public deficits and shrinking public resources, geographical targeting may be a viable way to allocate resources for poverty alleviation in developing countries. Efficiency can be increased and leakage to the nonpoor reduced substantially by targeting increasingly smaller areas. This article, and more generally the symposium on geographical targeting for poverty alleviation, proposes several techniques for augmenting data to produce more detailed poverty maps. It focuses on practical considerations in the design of geographically targeted poverty alleviation programs. In particular, it assesses the advantages and disadvantages of geographical targeting and describes how geographic information systems can be applied to improve poverty mapping.

Why have some geographic areas become pockets of poverty, while others have become islands of prosperity? Many explanations have been offered for the striking differences in living standards between regions and even between communities within the same region. Such disparities can be found in all countries, and they may be caused by a wide range of factors, including differences in agroclimatic conditions, endowments of natural resources, or geographic conditions—particularly the distance to a sea outlet and to centers of commerce—and biases in government policies.

Consider a few examples. Mean per capita consumption of the rural population in the Indian state of West Bengal is only half that of rural Punjab, and the headcount measure of poverty in West Bengal is nearly four times higher than that in Punjab (Datt and Ravallion 1993). In Burkina Faso the incidence of poverty is less than 25 percent in one-fifth of the villages, but well over 60 percent in more than half of the remaining villages (Bigman, Dercon, Guillaume, and Lambotte, this issue). And in Ecuador the incidence of poverty varies from less than 10 percent of the population in some districts to nearly 60 percent in others (Hentschel, Lanjouw, Lanjouw, and Poggi, this issue). Indeed, in many developing countries the differences in living standards between regions are often larger than the differences within regions. Policymakers who seek to design the most

David Bigman is with the International Service for National Agricultural Research, the Netherlands, and Hippolyte Fofack is with the Africa Region at the World Bank. David Bigman was with the Africa Region of the World Bank when this work was done. The authors' e-mail addresses are *d.bigman@cgiar.org* and *hfofack@ worldbank.org*.

cost-effective strategies for reducing poverty must consider the implications of these large income disparities between geographic areas.

Studies of income inequality and poverty generally take the approach of the individualistic, human capital model, which explains differences in income and consumption between people by looking at differences in individual and household characteristics. However, differences in standards of living between regions and communities are often far too large to be explained by differences in individual or household characteristics alone. Disparities in living standards may persist because of obstacles to internal migration, which in some countries are the result of deliberate government policies and in all countries are the result of economic, demographic, and cultural factors.

Migration between rural areas, for example, is constrained by the lack of free land for cultivation; rural-urban migration is an option available to only the fittest people, leaving behind the very young and the very old. Moreover, migration is costly and risky, and frequently individuals do not have the information they need to decide whether, when, and how to migrate. Even rural-urban migration, common as it is, often evolves as a gradual process in which one member of a household moves to the urban center in search of employment and is later followed by other members of the household. Further obstacles to migration that particularly affect the poor are large household size, poor health conditions, low levels of human capital, and, in some countries, the "feminization" of poverty (poverty rates are disproportionately high among women). Wealthier and better-educated individuals are less restricted in their ability to migrate, and, as they leave poor areas, the standard of living in those areas declines even further.[1]

Pockets of poverty persist for other reasons as well:

- The low quality of public services, particularly in education and health, further impedes the accumulation of human capital and thus earning capacity.
- The poor condition of rural infrastructure limits trade and retards local investment and growth (Binswanger, Khandker, and Rosenzweig 1993).
- The low level of social capital in poor communities slows the diffusion and adoption of new farm technologies, thus reducing farmers' earning capacity (Foster and Rosenzweig 1995).
- The distance from urban centers inhibits trade, specialization in production, and access to credit.

As a result households in poor areas are less likely to escape the individual and community predicaments that keep them poor.

There are pockets of poverty in both urban and rural areas. The relative ease of identifying these areas makes the place of residence a possible criterion to determine eligibility for poverty reduction programs (Baker and Grosh 1994). The argument holds that although some benefits inevitably leak to the nonpoor

1. This, however, should be captured by the individual characteristics rather than the spatial characteristics that explain interregional inequalities.

who reside in target areas, and although some of the poor who reside in nonpoor areas will not be covered, geographical targeting has several clear advantages. It is easy to implement and to monitor, thus it typically involves less fraud and much lower administrative costs than many other targeting methods. Moreover, in many developing countries the limited information available on individuals and households reduces the options for implementing other types of targeting. At the same time, severe budgetary restrictions make nontargeted programs infeasible.

The three articles in this symposium present different methods of identifying pockets of poverty in developing countries. The methods call for using all available information, even though it may be limited, in order to map the spatial distribution of poverty. Jesko Hentschel, Peter Lanjouw, Jean Olson Lanjouw, and Javier Poggi bring together household data from Ecuador's census and detailed data from its household income and expenditure survey in order to estimate the incidence of poverty at the village level and to provide a detailed profile of the spatial dimension of poverty in that country. David Bigman, Stephan Dercon, Dominique Guillaume, and Michel Lambotte use data from Burkina Faso's household survey; data from a wide variety of other surveys, including socioeconomic, health, and agricultural surveys; data from the country's population census; detailed climatic data; and detailed road-mapping data in order to determine the spatial distribution of poverty at the village level. These methods can be used in countries where census data provide only very limited information on household characteristics and where data from different surveys must be brought together and, if possible, compared in order to establish a reliable profile of the spatial distribution of poverty. In the third article Hippolyte Fofack uses Ghana's light monitoring survey together with the more comprehensive Integrated Survey to develop criteria for improving targeting.

The general principles of geographical targeting have been thoroughly researched and discussed in the economic literature.[2] This introduction offers a practical guide for using geographical targeting to place poverty alleviation programs and an overview of the questions that are central to selecting targeting criteria. In addition, it discusses applications of a geographic information system (GIS). In recent years the ability to incorporate geographic indicators in public policy planning in general, and in the design of welfare programs in particular, took a quantum leap with the development of new and sophisticated methods for incorporating spatial data and organizing these data as a GIS suitable for computer analysis. Equally important was the surge of technological innovations, such as satellite imagery, for collecting spatial and climatic data. By including in the GIS database not only information on social, economic, climatic, and environmental observations, but also their location and spatial arrangement, this system allows us to present the data in the form of maps and interfaces and to perform comprehensive and sophisticated spatial analysis. In many countries—both de-

2. See, for example, Kanbur (1987), Bigman (1987), Besley and Kanbur (1993), Ravallion (1993, 1998), Datt and Ravallion (1993), and Jalan and Ravallion (1998).

veloping and industrial—this system has become the single most important tool for analyzing a wide range of geographic and socioeconomic data and for designing policy measures.

I. Practical Considerations in Designing a Poverty Alleviation Program

Geographical targeting of welfare programs is common in developing countries and is often used in conjunction with additional targeting criteria to narrow the beneficiary population and thus reduce costs. For example, in Mexico the government gave food subsidies only in selected regions and only for tortillas and milk, the main staples of the poor. In Honduras the government restricted a food stamp program to selected areas and required means testing to determine eligibility. And in many Sub-Saharan African countries governments targeted the construction of new health and education facilities to relatively poor areas, and the services provided in these facilities primarily benefited the poor.

Informational Constraints

These and many other similar programs illustrate the intuitive appeal of geographical targeting, as well as its perils. The challenge for policymakers is to use the available resources to provide the greatest possible assistance to those who need it most. In the absence of reliable information on personal income, the first-best solution of identifying the poor and directing all benefits only to them is not feasible. Even in industrial countries that have the necessary data, it is not possible to ascertain whether targeted programs do indeed reach all of the poor and do not leak to the nonpoor.

In the past, since most developing countries did not have reliable information on individual income, many chose programs with universal coverage. In the 1960s and 1970s several countries in Sub-Saharan Africa and South Asia implemented general food subsidy programs. Growing budget constraints in the 1980s, however, forced governments to drastically reduce or even terminate these programs. Some countries replaced universal coverage with means testing, which initially proved quite successful. Sri Lanka is a typical example. In the late 1970s the cost of a universal ration program reached 5 percent of gross domestic product (GDP), and the government was forced to cut costs by replacing it with a food stamp program that cost only 1.3 percent of GDP. However, in the absence of reliable information on household income, means testing led to massive leakage: nearly half the population had access to food stamps, although less than 30 percent were eligible under the program's criteria (Subbarao and others 1997).

The absence of reliable information for identifying the poor, on the one hand, and the mounting constraints on public resources, on the other, made targeting by means of indirect indicators the only viable alternative for most developing countries. The indicators used to determine eligibility included the household's size, the number of children in the household, the size of the

household's landholdings or other assets, and the region in which the household was located. Other alternatives were self-targeted programs (such as food-for-work programs), subsidies for commodities consumed primarily by the poor, and research and extension services targeted to crops consumed primarily by the poor. But these programs were limited in duration and in coverage, thus excluding many of the poor and involving substantial leakage (Kanbur, Keen, and Tuomala 1994).

Targeted programs that use indirect criteria also entail high administrative costs and may have undesirable effects on the target population. For example, Lanjouw and Ravallion (1999) analyze participation rates in various social programs in India and find that it is the nonpoor who capture the early benefits. In family assistance programs implemented in several Latin American countries, eligibility was determined by the number of children in the household. The cost of these programs was pushed to an intolerably high level by slack entitlement conditions and high rates of leakage. Further, they raised the fertility rates of the poor. In Tanzania the difficulty of establishing clear eligibility criteria for distributing food aid forced the government to delegate the authority for distribution to specialized local nongovernmental organizations (NGOs) and village committees, which drove up costs.

Political Feasibility

The effectiveness of a targeted poverty alleviation program thus depends on the availability of efficient indirect criteria for identifying the poor, on the administrative costs of establishing eligibility with these criteria, on the reduction in costs that can be achieved by effectively excluding nonpoor households from the program, and on the government's capacity to administer the program. Grosh (1994) and Gelbach and Pritchett (1997) also emphasize political feasibility as one of the central factors determining a program's effectiveness and sustainability. The main obstacle to targeted programs has often been the opposition of population groups not covered by the programs, primarily middle- and high-income groups. Anand and Kanbur (1990) report that, after the Sri Lankan government introduced a targeted food stamp program, the real value of the food stamps fell sharply during periods of high inflation, as the interest of the middle class shifted to other issues, and public support for the program declined.

Political opposition to targeted programs may arise because only a portion of the general population is bearing their cost and because programs create a stigma for beneficiaries. Smolensky, Reilly, and Evenhouse (1995) distinguish between external and internal stigmas, that is, the stigma created because the welfare program lowers the self-esteem of participants and the stigma created by the society at large, which often leads to tensions between participants and the members of society who bear the tax burden of financing the program. Besley and Kanbur (1993) point out that by stigmatizing welfare recipients, income-based programs have reduced recipients' ability to acquire skills and grow out of poverty. Moffitt (1983) describes the stigma as the "disutility arising from participa-

tion in the welfare program"; Besley and Coate (1992) emphasize the "psychic costs of being on welfare."

In many cases, however, the most significant reason for the political opposition to a targeted program is leakage to ineligible households, which are seen to be free riding on the backs of taxpayers. Improved targeting can go a long way toward reducing such tensions. Rainwater (1982) points out, though, that more accurate targeting might also have the opposite effect of further stigmatizing the poor. This would occur particularly if eligibility required recipients to submit highly personal information (for example, the name of the father of a child born out of wedlock) in order to make entitlement more stringent. In some countries, particularly countries in Sub-Saharan Africa, targeted programs may also exacerbate ethnic tensions if the target group is perceived to be predominantly of a specific ethnic origin.

Advantages of Geographical Targeting

Geographical targeting offers several advantages over other methods of targeting. First, it provides clear criteria for identifying the target population and avoids the informational constraints that impede most other targeted programs. Second, it is relatively easy to monitor and administer, and local institutions and NGOs can greatly assist in implementing the programs. Third, geographical targeting has relatively little influence on a household's behavior, since it is difficult and costly for a household to change its place of residence. Fourth, it is possible to improve targeting by combining the geographic criterion with other eligibility criteria based on individual or household characteristics. Fifth, the instruments of geographically targeted programs can include not only direct income transfers to the target population but also a wide variety of other measures aimed at increasing the living standards of the entire population of the area. Examples include investment in infrastructure, provision of public health and education services, and provision of financial services. Geographical targeting thus can provide guidelines for allocating resources under a country's welfare program as well as under its development program.

The basic rationale for targeting poverty alleviation programs on the basis of geography is the existence of large differences in living standards between geographic areas and the concentration of poverty in some areas. These differences can be found in all countries: the western (inland) areas of China, parts of northeastern India, northwestern rural areas in Bangladesh, northern Nigeria, the rural savannah in Ghana, the northeastern region of Brazil, and the deep South in the United States are just a few examples of pockets of poverty. We even find large disparities in the standard of living between villages and urban communities within the same agroclimatic region. The shantytowns of Johannesburg, the *favelas* of Rio de Janeiro, and the slums of New York City exist side by side with affluent neighborhoods.

These disparities arise because of large differences in the price and quality of housing, the quality of physical infrastructure (primarily the quality of roads),

socioeconomic characteristics of the population, and the quality of public services (particularly health and education)—which perpetuate the cycle of poverty. Low housing costs in poor neighborhoods attract migrants from rural areas and deter local residents who manage to raise their incomes from staying, thereby deepening the pockets of poverty. The quality of the road infrastructure—primarily the availability of all-weather roads—is an important factor in determining an area's development and capacity to trade. And the concentration of welfare recipients in some areas stigmatizes these areas and deters private sector investment and trade.

Large disparities between regions and, even more so, between communities within regions, also arise because of differences in the geographic distribution of government spending, primarily spending on infrastructure (Hammer, Nabi, and Cercone 1995; van de Walle 1995). Such differences reflect the limited political power of poor areas, as well as the efforts of the government to concentrate investment in areas that have strong potential for growth. These two factors explain the urban bias of many governments in developing countries. In rural areas governments tend to invest in regions with good agricultural potential or with a concentration of natural resources, rather than in marginal lands, where the rural poor are concentrated. Election politics also account for large differences in the allocation of public funds, for example, between the relatively densely populated, mostly urban areas and the more sparsely populated, mostly rural areas. In some countries this bias also reflects ethnic, nationalistic, or religious differences between populations.

Although these large disparities and the presence of pockets of poverty make geography an attractive indicator for targeting poverty alleviation programs, the empirical issues and practical difficulties involved in the selection of target areas are far from settled. The main question, which has been examined in a large number of empirical studies, is whether geographical targeting is a cost-effective alternative to universal coverage or to other methods that use proxy indicators as substitutes for means testing.[3]

Baker and Grosh (1994) evaluate the potential impact on poverty of geographical targeting at different levels of aggregation, using household survey data from Venezuela, Mexico, and Jamaica. Their results indicate that geographical targeting can be a useful mechanism for transferring resources to the poor, and the reduction in poverty that can be achieved is larger than the reduction that can be achieved through an equally expensive universal distribution program. These results held for a general food subsidy program and for a food stamp program that used means testing as a self-selection process. Baker and Grosh also demonstrate that the level of geographic aggregation has a noticeable impact on the outcome of targeting—targeting smaller geographic areas makes it possible to improve efficiency and reduce poverty by a greater amount.

3. See Bigman (1987), Ravallion (1993, 1998), Datt and Ravallion (1993), Ravallion and Wodon (1997), Baker and Grosh (1994).

Ravallion and Wodon (1997) examine the significance of two sets of indicators in determining a household's well-being. One is based on the household's characteristics, and the other is based on the household's location. Using household data from Bangladesh, they show that a household's geographic profile is a more significant indicator of poverty than its other characteristics. Their results for Bangladesh indicate, however, that the gains from geographical targeting at the regional level are small. In a similar study using data from Indonesia, Ravallion (1993: 464) concludes, "The gains to the poor from geographical targeting at the level of regions could be quite small, even with large regional disparities in poverty."

Most of these studies evaluate the cost-effectiveness of geographical targeting when the target areas are administrative regions, provinces, or federal states. There are two reasons for concentrating on these levels. First, most studies use data from the country's household income and expenditure survey to evaluate a program's benefits to the poor and the nonpoor. The number of households selected for this survey is low, and the size of the sample from an area smaller than a region is too small to allow inferences that are statistically significant. Second, in most geographically targeted programs the target areas are the administrative region, the province, or the federal state. At these levels, however, income disparities within an area are, in most cases, still large, and target areas are bound to include many nonpoor households. Regional targeting therefore is likely to result in high leakage and, by excluding many, if not most, of the country's regions, it also leaves out a considerable portion of the country's poor.

Targeting at the Village or Community Level

Within smaller geographic areas, particularly rural areas, income disparities tend to be much smaller. This is because typically smaller areas have more homogeneous socioeconomic characteristics, and the population is subject to the same agroclimatic and geographic conditions. Targeting smaller administrative areas—districts, subdistricts, or even individual villages and urban neighborhoods—can therefore reduce leakage significantly. Moreover, if the program's targets are individual villages, then it is more likely that the selected villages will be spread across all regions, rather than concentrated in a few. This will reduce the ethnic and political tensions that may accompany regionally targeted programs because beneficiaries will include people from all ethnic or national origins and religions.

Further, by including the poorest districts or villages in all regions, targeting at these levels is likely to increase the proportion of the poor who are covered. Targeting smaller areas will also increase the choice of policy instruments that can be used to combat poverty. If a poverty alleviation program is targeted to villages, for example, it may provide a local source of drinking water in some villages, offer a food-for-work program in others, and construct an all-weather access road in still others. This issue will be discussed further below.

Even though targeting smaller geographic areas is an attractive strategy theoretically, it is usually applied only to very specific programs in health and education, not to a more general strategy for reducing poverty. Considerable practical

obstacles arise during implementation because it is difficult to obtain reliable estimates of the incidence of poverty in small areas and thus to determine eligibility. The articles in this symposium develop statistical and econometric methods for mapping poverty in small areas, including villages. These methods may go a long way toward overcoming these obstacles. To examine the practical difficulties and potential benefits that may arise in implementing poverty alleviation programs targeted to small areas, we focus here on programs targeted to individual villages or urban communities; very similar conclusions apply, however, to programs targeted to districts or subdistricts.

One potential difficulty with targeting programs to individual communities is selecting the criteria for determining eligibility. Unless the household survey contains reliable information on the incidence of poverty in each community in the country, we must use indirect indicators that are closely correlated with the incidence of poverty. Some indicators already are widely used in developing countries, often with strong incentives from international organizations or donor countries. Thus, for example, the combination of climatic conditions and the distance from urban centers often identifies villages in marginal lands whose standard of living is much lower than in the rest of the country. The quality of access roads and the distance to sources of drinking water are additional reliable and widely used indicators of living standards, primarily in rural areas. Still another group of indicators includes the availability of public services, the distance from the nearest public school or health clinic, and the type of services provided in the local health clinic. These indicators are obvious selection criteria in choosing the location of new health clinics or public schools; all too often they also indicate the incidence of poverty in the community, both because the relatively well-to-do tend to leave the community if services are inadequate and because poor services damage the earning capacity of the local population.

In the absence of reliable data for each village, however, these criteria are only approximations. They are rarely used to determine eligibility for more general social welfare programs, such as food stamps or child support, and they may generate considerable resistance. Not only will neighboring villages protest, but the urban population may also rail against the use of such criteria if the government does not make a convincing case that the selected villages are indeed likely to be the poorest and that the chosen criteria are highly reliable.

Another issue is the ability of village institutions to administer and monitor the program. At present, local authorities usually only take part in distributing benefits. However, if they are strong, local institutions, particularly religious institutions, can be very effective in fully administering general poverty alleviation programs at the village level. This may be less effective and more costly in villages where local leadership is weak. The main obstacle that village institutions will encounter is deciding who is a resident and therefore entitled to benefits. Disputes between neighboring villages may erupt. And the program may become less effective, since the marginal cases—the households that reside at the outskirts of the village—are often also the poorest and the most likely to be left out.

An important advantage of selecting villages or urban communities as targets is that it widens the choice of policy instruments. At present, different instruments are designed to advance different goals, such as increasing school enrollment or lowering child morbidity, rather than the same general goal of alleviating poverty. As a result, the education department—of the country or the World Bank—prepares a plan for constructing a new school in the community, taking into account only the availability and proximity of existing schools. Very little consideration is given in that plan to other needs of the village and other possible projects, such as improving the access road or health care services for children, that could be more important for improving the standard of living. Likewise, when an administrative department prepares a plan to improve the local road infrastructure, it takes into account the current state of access roads, but not the other needs of the village. Further, there is very little coordination among the departments responsible for these plans.

The need for coordination in designing public projects also arises when the target areas are regions. Suppose that a program for constructing new health facilities identifies target regions on the basis of regional poverty indicators. It is still necessary to determine the exact location of the facilities within the region. An obvious, and the most commonly used, criterion is the distance from existing health facilities to villages within the region. The second consideration is the cost of construction. Program designers often pay little attention to the standard of living in the village, the availability of other public services, and the potential of other projects to reduce poverty. Even distance to a facility can be a complex indicator if one takes into account access to the village during the rainy season, the available modes of transportation (which can be different in villages closer to urban centers and in more remote villages), and the type of health services required in each village.

The possibility of producing reliable poverty maps at the village level opens the door to a more comprehensive strategy for poverty alleviation that proceeds in the opposite direction. In the first stage the general goal of the program—be it the reduction of poverty or child mortality—is determined. In the second stage the policy instrument is tailored to the specific conditions of the local community, on the one hand, and the program's goal, on the other. Only in the third stage are the appropriate departments called on to design health clinics or rural roads—in accordance with that general plan. This approach, which considers the overall needs of the village, requires much more planning, coordination, and data. The data problems have been largely solved in the past decade thanks to significant advancements in the technology of the geographic information system.

II. The GIS and Its Applications

Space is central to the decisions, behavior, and even characteristics of individuals and communities. Nevertheless, until recent years it has received little

attention in the social sciences. In statistical and econometric analysis the use of spatial data was limited by the lack of an easy and effective way to incorporate spatial characteristics explicitly into models. The use of the GIS in the social sciences has progressed very slowly, and the application of spatial data analysis in econometric studies is still quite limited. The GIS is used primarily to display, organize, and run simple manipulations of spatial data.

A GIS is a computer-based system used to capture, store, edit, display, and plot geographically referenced data (Longley, Goodchild, and Maguire 1999). Mapping by means of a computer rather than the traditional cartography process can greatly reduce production time and costs. A GIS is also used, primarily in research, for integrating heterogeneous, geographically referenced data sets and providing a common reference framework for analying spatial data in different subjects and over different time periods. Data are integrated by using space as an indexing system. The spatial coordinates of the data make it possible, for example, to combine socioeconomic information on villages from household and community surveys with information on the surrounding farming systems from an agriculture survey and with information on climate and soil conditions from an agroclimatic study. By combining these data sets, a GIS enables us to conduct a comprehensive analysis of social and economic phenomena. In many countries this was the main incentive for compiling national data on infrastructure including basic data on roads, hydrology, settlements, and political and administrative boundaries.

The articles in this symposium present several methods of integrating spatial data in economic and social studies. Hentschel, Lanjouw, Lanjouw, and Poggi combine data from Ecuador's census and its household income and expenditure survey at subregional levels. Bigman, Dercon, Guillaume, and Lambotte integrate household-level data collected in Burkina Faso's household income and expenditure survey with village, district, and regional data collected in a wide variety of other surveys and in the population census. Fofack combines data from the Integrated Survey of Ghana, which has comprehensive information on household income, expenditures, and assets, with a light monitoring survey, which has large coverage and is exhaustive on the location and access to public facilities, in order to draw inferences for targeting.

A GIS allows a wide variety of data integration forms. One layer of data (such as districts) can be presented on top of another (such as climate zones), not only to create a visual display but also to generate a new data set in which each point has attributes from the two original data sets. The points in the integrated data set can then be used to analyze social, economic, and spatial relationships using either cross tabulations or formal statistical and econometric methods. Thus, for example, information on the distance from a village to an urban center can be combined with area-based soil data in order to assess the agricultural potential of rural communities; information on the road network can be combined with information on population density to generate indicators of transportation density for each district. Standard computations using GIS data include straight-line and

network distances, delineating the area that lies within a specified threshold distance from selected features or places. These functions can be used to determine, for example, the nearest hospital for each of a number of settlements or to identify all villages located within a certain distance from public schools.

III. Data Requirements

The data that are most commonly used for estimating poverty come from household income and expenditure surveys. The sample in these surveys is typically small, and the analysis must rely on additional data sources. In most developing countries the principal sources of information on all households with large coverage are the population and housing censuses. However, a census typically is conducted only once a decade, and the census questionnaire contains only limited information on the standard of living. Even so, the census remains one of the most important sources of information on demographic and social conditions.

Household surveys provide more detailed data, but they cover only a relatively small sample of households. Three examples of comprehensive household surveys in developing countries are the World Bank's Living Standards Measurement Study (LSMS) surveys, Integrated Surveys, and the U.S. Agency for International Development's (USAID's) demographic and health surveys. The LSMS surveys and Integrated Surveys concentrate on household income and consumption, while the demographic and health surveys focus on health indicators, including anthropometric measures that indicate the adequacy of food consumption and health care. Because the sample sizes of these surveys are relatively small, it is not possible to generate statistically reliable information for geographic areas smaller than a region. With the aid of a GIS, it is possible to aggregate the survey data according to geographic areas that are different from the administrative regions from which the sample of households was selected. For example, per capita expenditure indicators can be aggregated for agroclimatic zones, and health indicators can be aggregated for areas classified by their access to health services. Additional information from other surveys or from the population census organized as a GIS can then be used to determine, for example, the distance to the market town or the quality of the access road for each survey cluster.

In a rapidly changing economic and social environment, there are rapid and sometimes far-reaching changes in the incidence and spatial distribution of poverty. The surveys must be designed to provide up-to-date information. In most developing countries poverty is strongly affected by macroeconomic shocks, fluctuations in commodity prices, and the spread of diseases, particularly AIDS. These effects are not distributed evenly across economic regions, and, as a result, the geographic distribution of poverty may change considerably in short periods of time. A fall in commodity prices affects predominantly the areas in which these commodities are produced; macroeconomic shocks and a rise in unemployment affect predominantly urban areas, although rural areas also suffer from the re-

duction in government expenditures on rural infrastructure, health, education, and agricultural research; the AIDS epidemic is concentrated in certain geographic areas and affects a very large proportion of the population in these areas. In addition, rural-urban migration continuously deepens the concentration of the poor in urban areas.

These changes require frequent adjustment of poverty reduction instruments and more frequent collection of relevant data to produce an accurate mapping of poverty. In many countries, however, there has been a considerable reduction in the frequency of comprehensive household surveys. After a wave of comprehensive surveys in a large number of developing countries in the 1980s, only a few countries conducted further comprehensive surveys. Instead, many countries conducted far less comprehensive surveys, such as light monitoring surveys or the recent *enquêtes* 1, 2, and 3 that were carried out in several Sub-Saharan African countries and were restricted to large cities. In addition, the time between consecutive surveys has been continuously prolonged.

Two recent initiatives in Sub-Saharan Africa sought to bridge the data gap by using auxiliary data to estimate the geographic distribution of food insecurity and vulnerability. One initiative is USAID's Famine Early Warning System, and the other is the Food and Agriculture Organization's Food Insecurity and Vulnerability Mapping System. These two systems generate comprehensive spatial indicators that are related to the level of well-being and identify the geographic areas that are most vulnerable to shortfalls in food supply and to poverty. Despite the importance of this information, these two systems have only limited use in designing poverty alleviation programs. The main reason is that they describe the spatial distribution of vulnerability at the supranational level, whereas poverty alleviation programs are designed at the national level. However, poverty maps can be updated more regularly for geographically targeted poverty interventions using instrumented consumption variables obtained by combining light monitoring surveys, which are much cheaper and easier to implement, with more comprehensive household surveys (Fofack, this issue).

IV. Performance Criteria

The performance of a targeted program depends on the criteria that are used to determine eligibility for and desirability of the program, and on the instruments that are used to transfer benefits to the target population. A number of performance criteria have been suggested to evaluate the desirability of a program. They include the contribution to poverty reduction, the effect on household behavior, budgetary costs, and errors of inclusion (type I) and exclusion (type II). The error of inclusion measures the number of poor individuals who are excluded from the program, and the error of exclusion measures the number of nonpoor people who are included in the program.

Of these criteria, errors of inclusion and exclusion generally receive the greatest attention because of their intuitive appeal and direct budgetary implications.

Ravallion and Chao (1989) suggest a quantifiable performance measure for targeted programs that takes into account both of these errors. It defines the gains from targeting as the amount by which the budget for a nontargeted program would have to be increased in order to achieve the same reduction in poverty—measured by the poverty gap ratio—that can be attained through targeting. They term this measure the "equivalent gain from targeting." Clearly, the larger is the type I error, the higher are the costs of the targeted program, and the smaller is the equivalent gain. Likewise, the larger is the type II error, the smaller is the cost increase needed to provide the same reduction in poverty, and the smaller is the equivalent gain.

A corollary to this criterion is one that measures the reduction in the poverty gap that can be achieved with the targeted program compared with the reduction in the gap that can be achieved with the nontargeted progam, when the costs of the two programs are the same. The equivalent gain in this case would measure the ratio of the reduction in poverty with a targeted program relative to the reduction in poverty with a universal coverage program.

An alternative performance criterion compares the costs and effectiveness of the targeted program under consideration with the costs and effectiveness of another targeted program. We can term this measure the opportunity cost of targeting. In practice, this measure is more relevant than the equivalent gain, since the alternative of targeting a program to a specific region is to target it to another region, rather than to switch to a nontargeted program. It is easy to see, however, that the two criteria will rank target areas in the same order.

V. A Review of the Articles in this Issue

The articles in this issue present alternative methods of combining survey data and census data from a wide variety of sources and illustrate how to use the extended data sets to determine the spatial distribution of poverty and the effectiveness of alternative policy instruments. These methods can provide important tools for more comprehensive and coordinated poverty alleviation programs that focus on the needs and specific conditions in villages and urban neighborhoods.

Hentschel, Lanjouw, Lanjouw, and Poggi present a method for mapping poverty to guide the allocation of resources for public projects. This method combines household survey data on income and consumption, which are available for only a sample of communities and households, with census data, which are available for all communities and households, in order to estimate per capita consumption for all households included in the census. This article shows that poverty measures calculated with these consumption estimates closely match the measures calculated with the original household survey data. Although the consumption estimates are unbiased, their standard errors remain large, and they therefore provide only a first cut at drawing a detailed poverty map.

Bigman, Dercon, Guillaume, and Lambotte present another method of bringing together and analyzing data from different sources. They geo-reference the

data from all sources and bring them together at the village level on the basis of the names of the villages and the geographic coordinates identifying their location. They integrate agroclimatic and environmental data into the analysis at the provincial or regional level. Organizing the data as a GIS enables them to incorporate data on the distance from each village to other villages, to towns, and to public facilities such as schools and health clinics. Some of the data were available for all villages in the country; others, including data on household expenditures and health, were available only for a sample of households and villages.

The authors use the sample of villages included in the household survey to identify the significant community variables that best explain the average level of well-being and the prevalence of poverty in each village. They then use these variables to estimate the level of well-being of all villages in the country. They divide all of the villages in the country into four categories of well-being, ranging from the poorest to the least poor. These estimates thus identify not only the villages that should be the target of antipoverty programs, but also the villages that should be the target of cost-recovery programs.

Fofack presents a third method of using the data collected in household surveys to obtain criteria for geographical targeting. His method is based on the use of a light monitoring survey in which households are asked only about a small sample of their expenditure items. The light monitoring survey is unlike the full LSMS or Integrated Surveys, which are more comprehensive, typically covering dozens of household expenditure items. The use of the light monitoring survey may, however, lead to significant errors in allocating resources. The resulting cost and leakage may far outweigh the savings from using a light monitoring survey instead of a full household survey. Fofack presents a methodology for estimating total household expenditures from a light monitoring survey and demonstrates the reduction in the rate of mistargeting and leakage that can be achieved. The regional poverty predictors derived from this methodology are shown to be effective instruments for targeting poverty alleviation.

VI. Concluding Remarks

Geographical targeting can be a significant alternative for allocating resources to poverty alleviation in developing countries, especially in the face of shrinking public resources. With the accumulation of geographic data and improvements in GIS techniques, much greater efficiency can be achieved with this targeted scheme. The articles in this symposium present alternative econometric methods for mapping poverty at increasingly smaller geographic areas. Two of the articles present methods for targeting at the village level. At this level the budgetary costs of geographical targeting and the leakage to the nonpoor can be significantly reduced. The main reason for this greater efficiency is that the population tends to be relatively homogeneous and variations in per capita income between households in a village are low. In large geographic areas, in contrast, income variance is much higher, and the potential for greater leakage may increase the budgetary

costs and reduce program effectiveness. However, the lack of necessary data and administration needed to oversee poverty reduction programs in many developing countries makes geographic targeting cost-effective, even for larger geographic areas.

REFERENCES

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Anand, Sudhir, and Ravi Kanbur. 1990. "Public Policy and Basic Needs Provision: Interventions and Achievement in Sri Lanka." In Jean Drèze and Amartya Sen, eds., *The Political Economy of Hunger*. Vol. 3. Oxford: Clarendon Press.

Baker, Judy L., and Margaret Grosh. 1994. "Measuring the Effects of Geographical Targeting on Poverty Reduction." Living Standards Measurement Study Working Paper 99. World Bank, Washington, D.C.

Besley, Timothy, and Ravi Kanbur. 1993. "The Principles of Targeting." In Michael Lipton and Jacques van der Gaag, eds., *Including the Poor*. World Bank Regional and Sectoral Study. Washington, D.C.: World Bank.

Besley, Timothy, and Stephen Coate. 1992. "Understanding Welfare Stigma: Taxpayer Resentment and Statistical Discrimination." *Journal of Public Economics* 40(1):165–83.

Bigman, David. 1987. "Targeted Subsidy Programs under Instability: A Simulation and Illustration for Pakistan." *Journal of Policy Modeling* 9(3):483–501.

Binswanger, Hans P., Shahidur R. Khandker, and Mark R. Rosenzweig. 1993. "How Infrastructure and Financial Institutions Affect Agricultural Output and Investment in India." *Journal of Development Economics* 41(2):337–66.

Datt, Gustav, and Martin Ravallion. 1993. "Regional Disparities, Targeting, and Poverty in India." In Michael Lipton and Jacques van der Gaag, eds., *Including the Poor*. Washington, D.C.: World Bank.

Foster, Andrew, and Mark R. Rosenzweig. 1995. "Learning by Doing from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy* 103(6):1176–209.

Gelbach, J. B., and Lant H. Pritchett. 1997. "More for the Poor Is Less for the Poor: The Politics of Targeting." Policy Research Working Paper 1799. Development Research Group, World Bank, Washington, D.C. Processed.

Grosh, Margaret. 1994. *Administering Targeted Social Programs in Latin America: From Platitudes to Practice*. Washington, D.C.: World Bank.

Hammer, Jeffrey S., Ijaz Nabi, and James Cercone. 1995. "Distributional Effects of Social Sector Expenditures in Malaysia, 1974 to 1989." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence*. Baltimore, Md.: Johns Hopkins University Press.

Jalan, Jyotsna, and Martin Ravallion. 1998. "Are There Dynamic Gains from Poor-Area Development Program?" *Journal of Public Economics* 67(1):65–85.

Kanbur, Ravi. 1987. "Transfers, Targeting, and Poverty." *Economic Policy* 4(1):141–47.

Kanbur, Ravi, Michael Keen, and Matti Tuomala. 1994. "Labor Supply and Targeting in Poverty Alleviation Programs." *The World Bank Economic Review* 8(2):191–211.

Lanjouw, Peter, and Martin Ravallion. 1999. "Benefit Incidence, Public Spending Reforms, and the Timing of Program Capture." *The World Bank Economic Review* 13(2):257–73.

Longley, Paul A., Michael F. Goodchild, and David J. Maguire, eds. 1999. *Geographical Information Systems: Principles, Techniques, Applications, and Management.* 2d ed. New York: John Wiley.

Moffitt, Robert. 1983. "An Economic Model of Welfare Stigma." *American Economic Review* 73(5):1023–35.

Rainwater, Lee. 1982. "Stigma in Income-Tested Programs." In Irwin Garfinkel, ed., *Income-Tested Transfer Programs: The Case For and Against.* New York: Academic Press.

Ravallion, Martin. 1993. "Poverty Alleviation through Regional Targeting: A Case Study for Indonesia." In Karla Hoff, Avi Braverman, and Joseph E. Stiglitz, eds., *The Economics of Rural Organization.* Oxford: Oxford University Press.

———. 1998. "Poor Areas." In Aman Ullah and David E. A. Giles, eds., *The Handbook of Applied Economic Statistics.* New York: Marcel Dekkar.

Ravallion, Martin, and Kalvin Chao. 1989. "Targeted Policies for Poverty Alleviation under Imperfect Information: Algorithms and Applications." *Journal of Policy Modeling* 11(Summer):213–24.

Ravallion, Martin, and Quentin Wodon. 1997. "Poor Areas or Only Poor People?" Policy Research Working Paper 1798. Development Research Group, World Bank, Washington, D.C.

Smolensky, Eugene, Siobhán Reilly, and Eirik Evenhouse. 1995. "Should Public Assistance Be Targeted?" *Journal of Post Keynesian Economics* 18(1):3–28.

Subbarao, Kalanidhi, Aniruddha Bonnerjee, Jeanine Braithwaite, Soniya Carvalho, Kene Ezemenari, Carol Graham, and Alan Thompson. 1997. *Safety Net Programs and Poverty Reduction: Lessons from Cross-Country Experience.* Directions in Development. Washington, D.C.: World Bank.

van de Walle, Dominique. 1995. "The Distribution of Subsidies through Public Health Services in Indonesia, 1978–87." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence.* Baltimore, Md.: Johns Hopkins University Press.

THE WORLD BANK ECONOMIC REVIEW, VOL. 14, NO. 1: 147-65

Combining Census and Survey Data to Trace the Spatial Dimensions of Poverty: A Case Study of Ecuador

Jesko Hentschel, Jean Olson Lanjouw, Peter Lanjouw, and Javier Poggi

Poverty maps provide information on the spatial distribution of living standards. They are an important tool for policymakers, who rely on them to allocate transfers and inform policy design. Poverty maps are also an important tool for researchers, who use them to investigate the relationship between distribution within a country and growth or other economic, environmental, or social outcomes. A major impediment to the development of poverty maps has been that needed data on income or consumption typically are available only from relatively small surveys. Census data have the required sample size but generally do not have the required information. This article uses the case of Ecuador to demonstrate how sample survey data can be combined with census data to yield predicted poverty rates for the population covered by the census. These poverty rates are found to be precisely measured, even at fairly disaggregated levels. However, beyond a certain level of spatial disaggregation, standard errors rise rapidly.

Poverty maps provide a detailed description of the spatial distribution of poverty within a country. They can be extremely valuable to governments, nongovernmental organizations, and multilateral institutions that want to strengthen the impact that their spending has on poverty. For example, many developing countries use poverty maps to guide the division of resources among local agencies or administrations as a first step in reaching the poor.

Poverty maps also can be an important tool for research. Researchers typically look for the empirical relationship between poverty or inequality and indicators of development, such as economic growth, using a cross-country regression frame-

Jesko Hentschel is with the Poverty Division of the Poverty Reduction and Economic Management Network at the World Bank, Jean Olson Lanjouw is with Yale University and Free University, Amsterdam, Peter Lanjouw is with the Development Research Group at the World Bank and Free University, Amsterdam, and Javier Poggi is with Superintendencia Banco y Seguros, Lima. The e-mail addresses of the first three authors are *jhentschel@worldbank.org, planjouw@worldbank.org,* and *lanjouw@econ.yale.edu.* The authors are grateful to the Instituto Nacional de Estadística y Censos in Ecuador for access to the unit record data of the 1990 census. They benefited from discussions with Uwe Deichmann, Stefan Dercon, Joanna Gomulka, Vassilis Hajivassiliou, Jeffrey Hammer, Berk Ozler, and Martin Ravallion, and received useful comments from three anonymous referees, participants at a seminar at the Suntory and Toyota International Centres for Economics and Related Disciplines (STICERD) at the London School of Economics in May 1997, and participants at the World Bank workshop Geographical Targeting for Poverty Reduction and Rural Development at the World Bank in November 1997. The authors also thank Jim Shafer for helpful assistance.

work.[1] It is difficult, however, to control for the enormous heterogeneity across countries, and heterogeneity may mask true relationships. Further, there is limited country experience that we can use to understand the determinants and effects of the distribution of welfare. Moving toward more microeconomic studies—relying on the distributional variation across communities within a single country—may be a way forward.

However, we are hampered in developing poverty maps by the scarcity of disaggregated data. For example, indicators based on income or expenditures often are favored, but the information required for a finely disaggregated map based on income or expenditures is not generally available for a sufficient number of households. The World Bank's Living Standards Measurement Study (LSMS) surveys, which have been fielded in many developing countries, collect the information needed to construct comprehensive measures of income and consumption. But they are too small to allow for disaggregation beyond a simple rural-urban breakdown within broad regions of a country. Census data do not suffer from small-sample problems, but they typically contain limited information.

Many Latin American, African, and Asian countries have constructed welfare indexes designed to rank regions by combining basic census information, such as access to public services and level of education, and have used these indexes to build poverty maps.[2] The indexes sometimes are labeled "basic needs indicators." They generally are constructed in a fairly ad hoc manner and are restricted to the limited qualitative (and not quality-adjusted) data available in a census. Such indicators may be poor proxies for household consumption. Using detailed household survey data for Ecuador, we show that a crude basic needs indicator and a comprehensive consumption measure yield markedly different welfare rankings. We explore how census-based maps can be improved when using an income- or consumption-based indicator of welfare.

In some situations one may want to consider a notion of welfare that reflects not just access to resources but nonincome components as well. For example, in evaluating education programs, one might want to account for the intrinsic value of education in addition to its influence on income or consumption. An appropriate welfare indicator might give greater weight to education than it would receive implicitly from a consumption or income indicator. But if, for example, one wants to figure out how to compensate households for a general change in price levels, a welfare measure based on a fairly narrow measure of consumption might be more suitable.

In general, we can construct a basic needs indicator, choosing weights reflecting the contribution of each variable to total household income as well as any

1. Deininger and Squire (1996) compile a large international database for this purpose. Bruno, Ravallion, and Squire (1998) use this database to explore the relationship between economic growth and inequality. See also Alesina and Rodrik (1994) and Fields (1989).

2. For recent descriptions of the derivation of such maps, see World Bank (1996) for Ecuador, Government of El Salvador (1995) for El Salvador, and FONCODES (1995) for Peru. Other Latin American countries using such poverty maps include Colombia, Honduras, and Republica Bolivariana de Venezuela.

direct contribution to welfare not captured in income. We do not expect rankings based on such an indicator to correspond to those based on consumption. However, we do want a broader measure to also reflect actual consumption. In this article we demonstrate how this may be done more reliably at a disaggregated level. When one is concerned with a broader measure of welfare, these more accurate predictions of poverty based on consumption simply need to be combined in some manner with the other indicators considered relevant given the policy issue of concern.

Using an LSMS data set for Ecuador, we estimate models of consumption expenditures, restricting the set of explanatory variables to those that also are available in Ecuador's most recent census. We apply the parameter estimates from these models to the census data to predict the probability that a given household, taken from the census, is poor. We evaluate our approach by estimating the incidence of poverty in six broad regions and comparing these rates with rates estimated from the LSMS data alone.

We consider some of the statistical issues that arise because the poverty figures are predicted. Our approach yields unbiased estimates of the incidence of poverty from the census data, so that the expected prediction errors are zero. However, the poverty estimates do have standard errors, which must be calculated along with the poverty rates. These standard errors are small at levels of regional disaggregation likely to be of practical relevance, which is encouraging, but the errors become sizable when the poverty rates are calculated over very small groups of households. We thus are warned against disaggregating the poverty map excessively.

I. How Well Does a Basic Needs Indicator Target Poverty?

In this section we examine how effectively a basic needs indicator is able to identify the poor when the poverty indicator is consumption expenditures.[3] The basic needs indicator that we consider was developed in 1994 by the National Statistical Institute of Ecuador (Instituto Nacional de Estadística y Censos, INEC) in response to the government's request for a directory of poor households. The government was considering eliminating its gas subsidy and wanted the directory to target poor households for a compensatory transfer. It happens that the government did not remove the subsidy, and we do not want to imply that INEC regarded the basic needs indicator as anything other than a fairly crude measure

3. Consumption is an imperfect measure of the standard of living, but a comprehensive measure of consumption expenditures comes reasonably close to capturing a household's *achievement* of well-being in accordance with its own chosen bundle of goods and services. The choice between income and consumption also merits attention. For developing countries probably the most compelling argument in favor of consumption is that it typically is easier than income to measure accurately. Its relative smoothness across seasons or even years may make it a better indicator of long-term living standards than a measure of current income (see, however, Chaudhuri and Ravallion 1994). For more discussion see Atkinson (1989), Ravallion (1994), and Sen (1984). Hentschel and Lanjouw (1996) and Lanjouw and Lanjouw (1997) discuss the attraction of a comprehensive measure of consumption expenditures as an indicator of welfare.

Table 1. *Points Assigned to Services Included in* INEC*'s Basic Needs Indicator*

Level	*Water*[a]	*Sanitation*[b]	*Waste*[c]	*Education*[d]	*Crowding*[e]
1	100	100	100	100	100
2	50	50	50	50	75
3	25	25	25	25	50
4	0	0	0	0	25
5	n.a.	n.a.	n.a.	n.a.	0

n.a. Not applicable.

Note: INEC is the National Statistical Institute of Ecuador.

a. Level 1 = public network; 2 = water truck; 3 = well; 4 = other.

b. Level 1 = in house, flush; 2 = in house, no flush; 3 = shared; 4 = other.

c. Level 1 = collection by truck; 2 = burned or buried; 3 = discarded; 4 = other.

d. Level 1 = household head has tertiary education; 2 = secondary; 3 = primary or literate: 4 = none or unknown.

e. Level 1 = one person or fewer to a bedroom; 2 = between one and two; 3 = between two and three; 4 = between 3 and 4; 5 = more than four.

Source: INEC (1996).

developed to meet an urgent government request at short notice. However, INEC's approach to constructing its indicator does resemble one that many countries have followed and therefore provides a useful example.

INEC constructed its basic needs indicator at the household level. It consists of a weighted composite of five variables capturing access to water, access to sanitation services, access to waste disposal services, education (of the household head), and a crowding index (the number of people per bedroom).[4] Each service was assigned a certain number of points according to its availability and its type or level (table 1). INEC assigned these points based on its own judgment, not as the result of our empirical analysis. For each household the value of the basic needs indicator was taken simply as the sum of points across services. The lower the sum, the poorer the household.

Using data from a recent household survey, the Ecuador Encuesta sobre las Condiciones de Vida (ECV), we can judge how well the basic needs indicator is able to identify households that are poor, as measured by consumption expenditures. The ECV for 1994 is a nationally representative household survey modeled closely on the World Bank's LSMS surveys. It provides detailed information on a wide range of topics, including food consumption, nonfood consumption, labor activities, access to services such as education and health, agricultural practices, and entrepreneurial activities. The survey was fielded by the Servicio Ecuatoriana de Capacitación (SECAP) in Ecuador during June–September 1994. It covered more than 4,500 households; after cleaning the data and checking for consistency, 4,391 households were left.

The survey design incorporated both clustering and stratification on the basis of the country's three main agroclimatic zones and a rural-urban breakdown. It also oversampled Ecuador's two main cities, Quito and Guayaquil, and covered

4. In other poverty-mapping exercises INEC has experimented with a wider range of variables. In El Salvador the government has constructed a poverty map using 12 variables (Government of El Salvador 1995).

Table 2. *Poverty Incidence under Alternative Definitions of Welfare*

Region	*Per capita consumption*	*Basic needs indicator*
Costa	0.35	0.39
Urban	0.26	0.18
Rural	0.49	0.76
Sierra	0.33	0.28
Urban	0.22	0.04
Rural	0.43	0.50
Oriente	0.59	0.65
Urban	0.20	0.03
Rural	0.67	0.76
National	0.35	0.35
Urban	0.25	0.13
Rural	0.47	0.62

Note: Figures are based on two alternative welfare indicators applied to the ECV household survey data.
Source: Authors' calculations.

about 1,374 rural households in total. Household expansion factors were added to the data set so that inferences could be made about population aggregates. World Bank (1996) analyzes the ECV data set as part of its study of poverty in Ecuador. Hentschel and Lanjouw (1996) construct consumption totals for each household, and World Bank (1996) bases all comparisons of welfare across households on their totals. Expenditures also have been adjusted to take into account regional variation in the cost of living based on a Laspeyres food price index reflecting the consumption patterns of the poor.

We compare poverty by region and area using the basic needs and consumption indicators (table 2). Since no poverty line was developed specifically for the basic needs indicator, we must infer poverty rates. We do this by equating the national incidence of poverty as measured by the basic needs indicator with the headcount rate, which we obtain using per capita consumption and the consumption poverty line of 45,476 sucres per person per fortnight (approximately $1.50 per person per day) developed in World Bank (1996). Hence we are asking how the regional ranking of poverty differs when poverty is defined using these two different indicators, but holding constant the fraction of the population identified as poor. We distinguish only between rural and urban areas and among the country's three main agroclimatic zones.

At this level of aggregation we derive the same rankings from the two alternative definitions of welfare. But regional differences are much more accentuated using the basic needs indicator. With the basic needs indicator rural areas appear more poor and urban areas appear less poor than with the consumption indicator. Among the rural areas, Oriente and Costa are poorer than Sierra, and among the urban areas, Costa is poorest, followed by Sierra and then Oriente. As World Bank (1996) emphasizes, under the consumption criterion the rankings of rural and urban areas between Costa and Sierra are highly unstable and easily over-

turned, depending on where one draws the poverty line and if one chooses to work with a poverty measure other than the headcount ratio. The basic needs criterion gives the impression that differences in well-being across regions are unambiguous.

Finally, we compare the performance of the two indicators at the household level. We follow the design of the planned intervention by taking the bottom 20 percent of households as the intended beneficiaries. First, we compute the total number of households represented in the ECV data and calculate that just fewer than 450,000 represent 20 percent of all households. Next, we calculate the total number of points for each household according to INEC's basic needs criterion and select the 450,000 households that score lowest. Last, we calculate the percentage of beneficiary households that fall into each household per capita expenditure quintile (column 2 in table 3). Since the target group is the first quintile, the percentage of beneficiaries in the first quintile indicates how well the basic needs indicator performs. Also, if all households were to receive the same amount of money, the percentage of beneficiaries in the first quintile represents the percentage of resources that would reach the targeted group. Only 41 percent of households identified by the basic needs criterion as constituting the poorest quintile are, in fact, among the bottom 20 percent according to the consumption criterion. Thus the leakage from an allocation based on the basic needs criterion would be very high: 60 percent of resources would go to households that are not in the lowest quintile, with almost 10 percent going to the top two quintiles.[5]

II. Predicting Poverty

To give our weighting scheme a more analytical basis, we consider developing a poverty map by imputing household consumption levels using census data.[6] We can do this only if certain data requirements are met. A household survey, such as the ECV in Ecuador, must be available and should correspond roughly to the same period covered by the census. In addition, unit (household)-level census data must be available. We were fortunate to have been granted access to the 1990 census data for Ecuador, covering roughly 2 million households. Although the 1994 ECV data were collected four years after the census, the 1990–94 period was one of relatively slow growth and low inflation in Ecuador, so it is reasonable to assume relatively little change.

The underlying intention of the method we propose here is similar to that of small-area and synthetic estimation procedures used in demographic and area

5. The basic needs indicator might not perform so poorly if the targeting scheme were aimed at, say, only urban areas or if an alternative cutoff point in the distribution were used.

6. Although imputing missing observations within a sample is a common procedure (see, for example, Paulin and Ferraro 1994), imputing values from a combination of different data sets (out-of-sample imputation) is less usual. Bramley and Smart (1996) combine the British Family Expenditure Survey with census information to estimate local income distributions. However, Bramley and Smart do not have access to unit-level data from both sources. They derive local income distributions, not from predicted household income, but from estimates of mean income and distributions of household characteristics.

Table 3. *Distribution of Beneficiaries Identified Using the Basic Needs Criterion and Predicted Consumption across Actual Consumption Expenditure Quintiles*

Quintile (as measured by per capita consumption)	*Percentage of beneficiary households (based on a basic needs indicator)*	*Percentage of beneficiary households (based on predicted consumption)*	
		Within-sample[a]	*Outside-sample*[b]
1 (poorest)	41.4	59.9	51.0
2	29.5	22.0	27.0
3	19.5	13.8	13.1
4	8.0	3.9	8.0
5	1.6	0.2	0.9

Note: Beneficiary households are those in the poorest quintile as measured by the basic needs indicator or by predicted consumption.

a. The within-sample exercise derives predicted household consumption from models estimated using the full household survey, applying the parameter estimates again to the full sample.

b. The outside-sample exercise consists of estimating the models for a subsample of the ECV and using the resulting parameter estimates to predict consumption for the remaining sample.

Source: Authors' calculations.

statistics.[7] In those procedures the interest is in deriving (unobserved) local-area attributes, such as a mean or total, often in the form of proportions (Farrell, MacGibbon, and Tomberlin 1997). For example, if we know population changes for a large area, we can use small-area estimation techniques to calculate population changes at lower geographic levels based on postulated functional relationships. The method we use here works in the opposite direction: we predict our variable of interest (consumption) at the unit (household) level and base aggregate statistics on those predictions.[8]

Estimating Models of Consumption

To impute expenditures using census data, we must first estimate a model of consumption using household survey data. Of course, the only variables that we can use to predict consumption are those that also are available in the census. In the case of Ecuador this set consists of demographic variables, such as the household's size and its age and sex composition; the education and occupation of each family member; the quality of housing (materials, size); access to public services, such as electricity and water; principal language spoken in the household; and location of the household. (See the appendix for comparative summary statistics from the two data sets.) The total number of explanatory variables, including dummy variables, interaction terms, and higher-order terms, is 48.

7. See Purcell and Kish (1980) for an overview. Isaki (1990) uses small-area estimation to obtain economic statistics.

8. The issue of combining information from different data sets has sparked a recent interest in the literature (see, for example, Arellano and Meghir 1992; Angrist and Krueger 1992; Lusardi 1996; and Imbens and Hellerstein 1999). These studies generally combine several household surveys, rather than census and survey data, and do not address spatial poverty estimation.

We estimate separate models for each region (Costa, Sierra, and Oriente), and within each region we distinguish between urban and rural areas. We also obtain separate estimates for Guayaquil and Quito, since the ECV oversampled these two cities.[9] The dependent variable in each regression is the logarithm of per capita consumption expenditures for household i, $\ln y_i$:

$$\ln y_i = \mathbf{X}_i'\boldsymbol{\beta} + \varepsilon_i \tag{1}$$

where $\mathbf{X}_i$ is a vector of independent variables common to the ECV and the census, and ε_i is a random disturbance term that is distributed i.i.d. $N(0, \sigma^2)$. We estimate the models with weighted least squares, using household sampling weights. The explanatory power of the rural models ranges from an R^2 of 0.46 for rural Sierra to an R^2 of 0.74 for rural Oriente. The explanatory power of the urban models ranges from an R^2 of 0.55 for Quito to an R^2 of 0.64 for urban Sierra.

We do not report here full sets of parameter estimates, standard errors, and diagnostics from the eight regression models for reasons of space, but these are available on request. It is important to correctly specify the precise functional form of the disturbance term in the consumption regression when calculating the second-stage poverty estimates. Thus we test the normality assumption made in equation 1. In three of the eight regions we cannot reject normality based on Shapiro-Wilk and joint skewness and kurtosis tests (all p-values > 0.15). Closer inspection of the residuals reveals that, in the other regions, we can reject normality only because of a few outliers in one or both of the tails. The outliers may be the result of mismeasurement. (For example, in one case the highest value of consumption expenditures is six times larger than the second-highest value.) After eliminating these observations—a total of only 13 out of 4,365—we cannot reject normality at conventional levels of significance in any region. Such small deviations from the assumed normality of the disturbance term should have a negligible effect on the accuracy of our results that follow. Further, with the exception of Guayaquil, we cannot reject (at the 10 percent level) the null hypothesis of homoskedasticity against the alternative of heteroskedasticity for the full set of independent variables.

9. It is of interest to consider how the basic needs weights in table 1 compare to those implied by the regression coefficients. Quito is a typical example. In the basic needs classification a decrease from four to three people per bedroom is associated with a welfare improvement equivalent to that from an increase in the education of the household head from primary to secondary school. The point estimates from the consumption regression suggest that an increase in the education of the household head from primary to secondary school is associated with an increase in consumption of 30 percent, while a reduction from four to three people per bedroom is associated with an increase in consumption of just 6.7 percent. An increase in education from secondary to tertiary also is associated with an increase in consumption of 30 percent. However, a decrease from three people to one person per bedroom, which has an equivalent welfare effect as moving from a secondary to a university education according to basic needs weights, is associated with a much larger increase in consumption: 47.6 percent. The same pattern holds across regions. Thus at high levels of crowding and low levels of education, the basic needs system gives more weight to reductions in crowding than to increases in education than would be appropriate given their relationship to consumption. If the basic needs weights are intended to reflect the relationships of both variables to overall consumption as well as an adjustment for their intrinsic value, then the weights seem to suggest a value judgment that being literate or attending primary school is less important than reducing crowded bedrooms.

Before moving to the second step, in which we apply the models to the census data, we test to see if predicting consumption (on the basis of the survey) improves targeting relative to the basic needs indicator. Although we obtain reasonable fits for cross-sectional regressions (as reported above), the coefficients of determination remain significantly lower than 1. To assess the performance of the model, we compare the basic needs indicator to actual consumption, as in the exercise reported in column 2 of table 3. We find that prediction models are better at identifying the poorest households—poorest in terms of consumption expenditures—than the basic needs indicator (columns 3 and 4 of table 3). In the first test we use the full household sample in the prediction models and apply the parameter estimates to the full sample (column 3). Targeting efficiency improves by almost 50 percent, with almost 60 percent of the bottom quintile as designated by predicted consumption also found in the bottom quintile as designated by actual consumption. The second test is considerably more demanding (column 4). Here, we (randomly) split the household survey in half and estimate the model of consumption using only half of the survey data. We then predict consumption for the other half of the sample (an out-of-sample prediction). As expected, the improvement over the basic needs indicator is less dramatic with this test. Nevertheless, if our goal is to target the bottom 20 percent of the population, this approach still improves targeting efficiency from 41.4 percent (basic needs) to 51.0 percent.

Predicting Poverty

We now proceed to the second step of the imputation exercise and apply to the census data the parameter estimates from the regressions (using the full household sample). For each household in the census we multiply its characteristics by the parameter estimates from the applicable regression (determined by the location of residence) in order to obtain an imputed value for the log of per capita consumption expenditures. We then estimate the household's probability of being poor, taking into account that the model does not perfectly explain consumption (the R^2 values never equal 1) and that predicted consumption is based on sample data. Finally, we calculate the incidence of poverty as the mean of the household-specific estimates for the population in a given region of the census.[10]

More formally, given a poverty line, z, the indicator of poverty, P_i, for each household i is

$$P_i = 1 \text{ if } \ln y_i < \ln z;\ P_i = 0 \text{ otherwise.} \tag{2}$$

10. Our discussion relies on a single poverty criterion—the incidence of poverty—and a single poverty line. One could, however, rank regions according to a range of poverty or inequality measures and experiment with a range of poverty lines (see Elbers, Lanjouw, and Lanjouw 2000). Also, our study examines the incidence of poverty among households. To calculate the incidence of poverty at the level of individuals, it is necessary to weight each household-level observation by the corresponding household size. The poverty figures provided in the tables are such weighted totals.

Using the model of consumption in equation 1, the expected poverty of household i with observable characteristics $\mathbf{X}_i$ is

$$E\left(P_i \mid \mathbf{X}_i', \boldsymbol{\beta}, \sigma\right) = \Phi\left[\frac{\ln z - \mathbf{X}_i'\boldsymbol{\beta}}{\sigma}\right] \tag{3}$$

where Φ is the cumulative standard normal distribution. Given that we are dealing with the headcount poverty indicator (equation 2), equation 3 is simply the probability that a household with observable characteristics $\mathbf{X}_i$ is poor.[11] From equation 1 we obtain estimates of $\hat{\boldsymbol{\beta}}$, the vector of coefficients, and $\hat{\sigma}$. Thus our estimator of the expected poverty of household i in the census is

$$P_i^* = E\left(P_i \mid \mathbf{X}' =_i, \hat{\boldsymbol{\beta}}, \hat{\sigma}\right) = \Phi\left[\frac{\ln z - \mathbf{X}_i'\hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right], \tag{4}$$

which, as a continuous function of consistent estimators is, itself, a consistent estimator of $E(P_i)$. P, regional poverty, is

$$P = \frac{1}{N}\sum_{i=1}^{N} P_i, \tag{5}$$

where N is the number of households in the region, and expected poverty is

$$E(P \mid \mathbf{X}, \boldsymbol{\beta}, \sigma) = \frac{1}{N}\sum_{i=1}^{N} E(P_i \mid \mathbf{X}_i, \boldsymbol{\beta}, \sigma). \tag{6}$$

The predicted incidence of poverty, P^*, given the estimated model of consumption, is thus

$$P^* = E(P \mid \mathbf{X}, \hat{\boldsymbol{\beta}}, \hat{\sigma}) = \frac{1}{N}\sum_{i=1}^{N} \Phi\left[\frac{\ln z - \mathbf{X}_i'\hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right]. \tag{7}$$

We calculate the incidence of poverty as the mean of households' probability of being poor rather than simply count households whose predicted expenditures are below the poverty line. The latter approach would give biased estimates of poverty rates.[12] Because of the random component of consumption, ε, no household has a zero probability of being poor or nonpoor, given its observed characteristics.

11. That is, if we were to take infinite draws from a population of households, the resulting poverty rate among households with observable characteristics $\mathbf{X}_i$ would be that given in equation 3. This value is not, in general, the same as the actual poverty rate, which is a *sample* from this infinite population, and depends on the particular realizations of ε_i.

12. This problem has been noted in the context of measuring the welfare of individuals, in which the bias arises because of inequality in intrahousehold distribution (Haddad and Kanbur 1990). See also Ravallion (1988). The Peruvian statistical institute—INEI (1996)—develops a model very similar to the one used here but derives poverty rates by directly estimating the headcount rate, not the predicted probability of being poor.

For each geographic region we compare the estimated incidence of poverty from the census data, using our imputed consumption values, with the rates obtained from the ECV household survey, using the consumption figures actually in the data (table 4). The incidence of poverty estimated from the ECV data in Ecuador as a whole is 35 percent. In general, poverty rates in the survey are reasonably close to, although somewhat lower than, those from the census (except in rural Oriente, for which the figures are the same in the two data sources). The differences are likely a result of changes in the exogenous variables underpinning the consumption regressions between the 1990 census and the 1994 ECV survey. For example, reductions in poverty are most apparent for Sierra, the region in which mean years of schooling of the household head appear to have risen most sharply between 1990 and 1994 (see the appendix). At the regional level, standard errors on the poverty rates calculated from the census are remarkably low.[13]

The two data sources do not rank the eight regions identically, but both clearly identify rural areas as poorer than urban areas, with rural Oriente emerging as the poorest region. World Bank (1996) indicates that orderings of regions, based on the ECV data, generally are not robust in that alternative poverty lines and poverty rates produce different rankings. The only exception is the rural versus urban ranking, which is found to be highly robust (first-order stochastic dominance held, with rural Ecuador being consistently poorer than urban Ecuador). The comparison of regional rankings based on the ECV and census data is consistent with these dominance results.

Standard errors on the ECV poverty rates (table 4) are such that we cannot reject the hypothesis that within sectors (urban and rural) poverty rates across regions are the same (although we can distinguish statistically between urban

13. For poverty incidence calculated from census data, the standard error of our indicator around the true poverty rate can be calculated as follows (see Elbers, Lanjouw, and Lanjouw 2000 for details):

$$P^* = \sum_{i=1}^{N} \frac{m_i}{M}\left(\frac{\ln z - \mathbf{X}_i'\hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right),$$

$$Var(P^*) \approx \left(\frac{\partial P^*}{\partial \hat{\boldsymbol{\beta}}}\right)' Var\left(\hat{\boldsymbol{\beta}}\right)\frac{\partial P^*}{\partial \hat{\boldsymbol{\beta}}} + \left(\frac{\partial P^*}{\partial \hat{\sigma}^2}\right)^2 \frac{2\hat{\sigma}^4}{n-k-1} + \sum_{i=1}^{N} \frac{m_i^2 P_i^*(1-P_i^*)}{M^2}$$

where n is the sample size for the consumption model with k parameters, estimated using the ECV survey; N is the number of households in the census population in the region of interest; m_i is the number of individuals in household i; and M is the total number of individuals in the census population.

$$\frac{\partial P^*}{\partial \hat{\boldsymbol{\beta}}_j} = \sum_{i=1}^{N} \frac{m_i}{M}\left(\frac{-x_{ij}}{\hat{\sigma}}\right)\phi\left(\frac{\ln z - \mathbf{X}_i'\hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right),$$

where ij indicates the jth element of the vector of explanatory variables for the ith household, and

$$\frac{\partial P^*}{\partial \hat{\sigma}^2} = -\frac{1}{2}\sum_{i=1}^{N} \frac{m_i}{M}\left(\frac{\ln z - \mathbf{X}_i'\boldsymbol{\beta}}{\hat{\sigma}^3}\right)\phi\left(\frac{\ln z - \mathbf{X}_i'\hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right).$$

Table 4. *Regional Poverty Rates for Ecuador*

Region	*ECV data*	*Census data*
Rural Oriente	0.67	0.67
	(0.02)	(0.004)
Rural Costa	0.50	0.52
	(0.02)	(0.002)
Rural Sierra	0.43	0.53
	(0.02)	(0.001)
Guayaquil	0.29	0.35
	(0.02)	(0.002)
Quito	0.25	0.33
	(0.02)	(0.002)
Urban Costa	0.25	0.29
	(0.02)	(0.002)
Urban Oriente	0.20	0.25
	(0.02)	(0.009)
Urban Sierra	0.19	0.29
	(0.02)	(0.003)

Note: Estimated standard errors are in parentheses.

a. Rates from the census are calculated using imputed expenditures based on a model calibrated from the ECV survey.

Source: Authors' calculations.

and rural sectors). Our estimates from the census data are sufficiently precise to permit meaningful comparisons across regions within sectors.[14]

III. Tracking Poverty at the Provincial Level

Using the methodology we have outlined, we can construct a poverty map, based on consumption expenditures, at a level of disaggregation below the eight broad regions for which the ECV is suitable. For example, there are nearly 400 cantons in Ecuador, each with some degree of local autonomy and administration, and these cantons can themselves be divided into more than 1,000 *parroquias* (parishes). Working with the census data, we can easily calculate expected poverty rates at the canton or parish level to determine where poverty is concentrated. In fact, as we have seen in the example described in section I, we can, in principle, use the census data to identify poor households and to target transfers to these households directly.

However, the standard errors on poverty estimates are a function of the degree of disaggregation of the poverty map (see the final term in the third equation of note 13). This warns us against attempting to use our methodology to identify, say, individual households that are poor.[15] Moreover, these objections come in addi-

14. Because the eight regions that we are comparing are based on different regression models in the ECV, the parameter estimates underlying the predicted expenditures are independent across regions. Therefore we can test for statistical significance of the difference in poverty rates between region r and region u based on the formula:

$$Var(P_r^* - P_u^*) = Var(P_r^*) + Var(P_u^*).$$

15. Suppose that the predicted probability of poverty for a given household is 48 percent. For a single household a lower-bound estimate of the standard error on that household's poverty rate would be $0.49 \cong \sqrt{0.48(1-0.48)}$.

tion to the well-known arguments against targeting in this way, which focus on the impact that such policies could have on the behavior of potential beneficiaries.[16]

Despite the dangers of micro-targeting, it may be desirable to develop a poverty map that is more disaggregated than broad regions. Ultimately, the optimal degree of disaggregation will depend on a number of factors. One is the precise purpose of the poverty map. Will it, for example, be used to identify government administrative areas so that the desired level of disaggregation is some level of local government? Or will it be used to identify poor villages or neighborhoods so that community-level projects (such as public infrastructure projects) can be better targeted? A second important consideration is whether we can assume that the parameter estimates from a regression model estimated, say, at the regional level, apply at subregional levels. Throughout this exercise we implicitly assume that, within a region, the model of consumption is the same for all households irrespective of the province, county, or community in which they reside. We cannot test this assumption, and at very fine levels of disaggregation it might be less appealing. (See Elbers, Lanjouw, and Lanjouw 2000 for a discussion of the implications of spatial autocorrections).

The desired degree of disaggregation also will depend on the availability of other sources of information, possibly local sources, on the poverty of individuals. Finally, other methods of local targeting, such as self-targeting, will become more important and effective at certain levels of disaggregation. Constructing a poverty map thus is likely to be a sequential process of gradual disaggregation until one reaches the point at which it seems there is no further insight to be gained.

Breaking down the headcount poverty rate by province, we see that poverty rates vary considerably (table 5). We also see that the standard errors on the poverty rates remain low, so that disaggregating to the level of provinces has not come at a significant cost in terms of statistical precision. A poverty map would have to be highly disaggregated before the standard errors would increase significantly because of small populations. In fact, only when the parish population falls well below 500 households do the corresponding standard errors rise enough to compromise comparisons (figure 1).[17]

16. Van de Walle and Nead (1995) provide a clear and thorough discussion of these issues.

17. We would calculate the standard error on the difference in poverty rates between two parishes in different regions as described earlier. However, because the parameter estimates determining the imputed expenditure figures are the same for all parishes within a given region, the standard error on the difference between two parishes in a given region is:

$$Var(P_1^* - P_2^*) = \left(\frac{\partial(P_1^* - P_2^*)}{\partial\hat{\boldsymbol{\beta}}}\right)' Var(\hat{\boldsymbol{\beta}})\left(\frac{\partial(P_1^* - P_2^*)}{\partial\hat{\boldsymbol{\beta}}}\right) + \left(\frac{\partial(P_1^* - P_2^*)}{\partial\hat{\sigma}^2}\right)^2 \frac{2\hat{\sigma}^4}{n-k-1} + \sum_{i=1}^{N_1} \frac{m_i^2 P_{1i}^*(1-P_{1i}^*)}{M_1^2} + \sum_{k=1}^{N_2} \frac{m_k^2 P_{2k}^*(1-P_{2k}^*)}{M_2^2}.$$

$$\frac{\partial(P_1^* - P_2^*)}{\partial\hat{\boldsymbol{\beta}}_j} = \sum_{i=1}^{N_1} \frac{m_i}{M_1}\left(\frac{-x_{ij}}{\hat{\sigma}}\right)\phi\left(\frac{\ln z - \mathbf{X}_i'\hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right) - \sum_{k=1}^{N_2} \frac{m_k}{M_2}\left(\frac{-x_{kj}}{\hat{\sigma}}\right)\phi\left(\frac{\ln z - \mathbf{X}_k'\hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right)$$

where N, M, and m are defined as in footnote 13 for parishes 1 and 2, which are subscripted by i and k, respectively, the subscript j indicates the jth element of the given vector, and

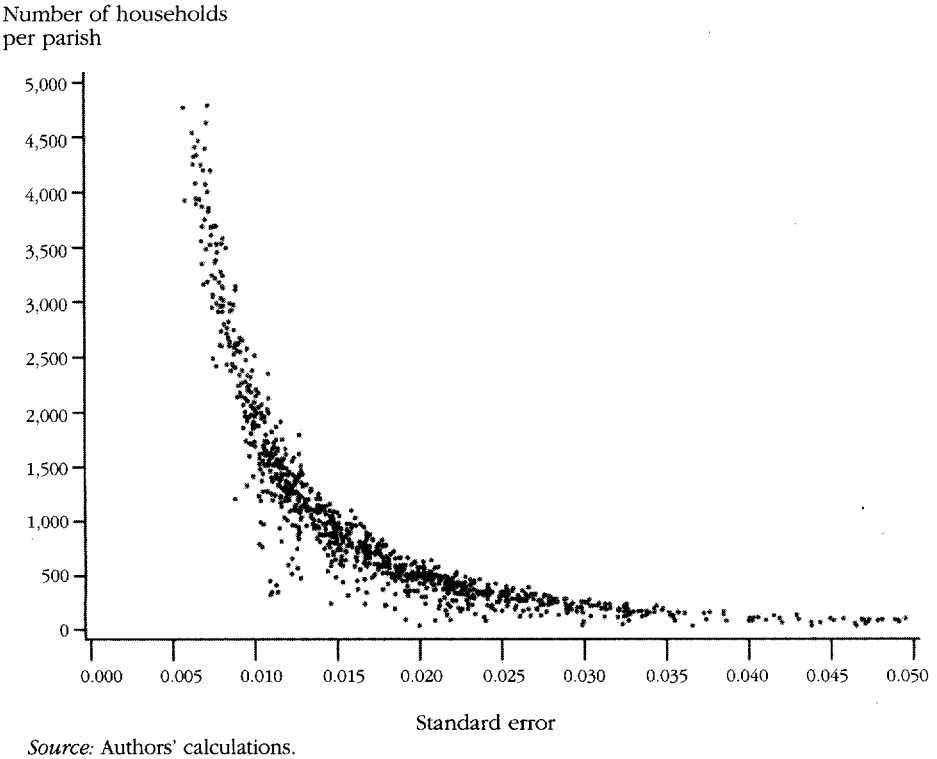
$$\frac{\partial(P_1^* - P_2^*)}{\partial\hat{\sigma}^2} = -\frac{1}{2}\sum_{i=1}^{N_1} \frac{m_i}{M_1}\left(\frac{\ln z - \mathbf{X}_i'\hat{\boldsymbol{\beta}}}{\hat{\sigma}^3}\right)\phi\left(\frac{\ln z - \mathbf{X}_i'\hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right) + \frac{1}{2}\sum_{k=1}^{N_2} \frac{m_k}{M_2}\left(\frac{\ln z - \mathbf{X}_k'\hat{\boldsymbol{\beta}}}{\hat{\sigma}^3}\right)\phi\left(\frac{\ln z - \mathbf{X}_k'\hat{\boldsymbol{\beta}}}{\hat{\sigma}}\right).$$

Table 5. *Poverty Map of Ecuador*

Region and province	*Expected poverty rate*	*Standard error*
Rural Oriente	0.67	0.004
Pastaza	0.65	0.005
Sucumbios	0.65	0.005
Morona Santiago	0.66	0.005
Zamora Chinchipe	0.67	0.005
Napo	0.69	0.004
Rural Sierra	0.53	0.001
Tungurahua	0.45	0.002
Pichincha	0.46	0.002
Azuay	0.50	0.002
Canar	0.52	0.003
Bolivar	0.55	0.003
Imbabura	0.56	0.003
Loja	0.57	0.003
Carchi	0.58	0.004
Chimborazo	0.59	0.003
Cotopaxi	0.63	0.003
Rural Costa	0.52	0.002
El Oro	0.45	0.003
Guayas	0.48	0.002
Los Rios	0.55	0.002
Manabi	0.56	0.002
Esmeraldas	0.59	0.003
Galapagos	0.14	0.008
Urban Costa		
Guayaquil	0.35	0.002
Urban Sierra		
Quito	0.33	0.002
Costa other urban	0.29	0.002
El Oro	0.24	0.003
Esmeraldas	0.27	0.004
Manabi	0.29	0.003
Guayas	0.30	0.003
Los Rios	0.32	0.003
Sierra other urban	0.29	0.003
Azuay	0.23	0.003
Tungurahua	0.25	0.004
Chimborazo	0.25	0.004
Cotopaxi	0.28	0.004
Loja	0.31	0.004
Canar	0.31	0.006
Imbabura	0.33	0.004
Carchi	0.33	0.005
Pichincha	0.33	0.003
Urban Oriente	0.25	0.009
Pastaza	0.24	0.011
Zamora Chinchipe	0.24	0.013
Morona Santiago	0.28	0.013

Source: Authors' calculations.

Figure 1. *Standard Errors on Headcount Rates and Population Disaggregation in Ecuador, Parish-Level Estimates*



Source: Authors' calculations.

IV. Concluding Remarks

In many developing countries poverty maps play an important role in guiding the allocation of public spending for alleviating poverty. A poverty map is essentially a geographic profile of poverty, indicating in which parts of a country poverty is concentrated and thus in which locations policies might be expected to have the greatest impact on poverty. A poverty map is most useful if it can be constructed at a fine level of geographic disaggregation.

To achieve such fine levels of disaggregation, it is essential to work with very large data sets. However, it is rare to find survey data that are both large in sample size and detailed in terms of household welfare. In general, there is a trade-off between size and quality because both goals are costly in financial and administrative terms.

In this article we have explored the possibility of combining the best parts of two different sources of data in order to construct a disaggregated poverty map that is based on an income or consumption measure of welfare. Constructing a poverty map based on census data, but using an ad hoc weighting scheme, may not be a good way to target those households deemed poor on the basis of their

consumption. Transfer programs based on such a map might reach only a subset of the intended beneficiaries and might entail considerable leakage to the nonpoor.

We suggest an alternative approach. Using household data from a high-quality, but small, living standards survey for Ecuador (SECAP 1994), we directly model consumption as a function of explanatory variables that also are present in the census. Because even the relatively few explanatory variables common to both the census and the ECV explain much of the variation in household consumption in the ECV, the incidence of poverty calculated from the census, based on this imputed consumption figure, is close to that calculated from the ECV. In Ecuador the poverty rates derived in the census generally are calculated with a high level of statistical precision. This precision declines as the degree of spatial disaggregation increases: although one might be tempted to use the methodology developed here to identify individual poor households, we demonstrate that such an application would be inappropriate. Our approach can be used at a high level of disaggregation but should be supplemented with complementary sources of information.

The most useful practical application of this methodology is probably in making comparisons with regional patterns of other indicators of well-being, opportunity, and access. For example, one could overlay our poverty map on a map documenting regional patterns of access to primary health care centers. Such an exercise could help policymakers decide where to prioritize efforts to expand access to primary health centers. It also could help policymakers decide how to expand such access—they might want to subsidize access in poor areas but experiment with cost-recovery methods in less-needy areas. Furthermore, a close correlation between, say, regional patterns of rural poverty and road access also might offer clues as to possible causes of poverty. This type of exercise could be undertaken for a wide range of indicators: levels of health and education, ethnicity and indigeneity, access to infrastructure and other public services, land quality and ecology, environmental conditions, and so on.

Finally, the ability to construct finely disaggregated poverty maps also might inform broader research questions. One direction would be to examine how the relationship between distributional outcomes and economic performance varies spatially within a country, in a manner analogous to cross-country analysis. This approach may avoid some of the methodological concerns arising with cross-country analysis. Other research questions also could be tackled. For example, underlying some of the current arguments in favor of decentralizing poverty programs is a notion that local communities themselves are best placed to identify the kinds of interventions that would be most beneficial to the poor. This position hinges somewhat on the contention that at the community level a subset of nonpoor households is less likely to capture public resources. This assumption probably is linked to the degree of inequality at the community level, something that traditionally has been difficult to measure. With the methodology presented here, household consumption inferred from the census could be analyzed to assess the extent of inequality within the community.

Appendix. *Comparative Descriptive Statistics from the 1994 LSMS and the 1990 Census*

	Rural Sierra		*Urban Sierra*		*Quito*		*Rural Costa*	
Indicator	*LSMS*	*Census*	*LSMS*	*Census*	*LSMS*	*Census*	*LSMS*	*Census*
Years of schooling of household head	4.48	4.33	8.75	8.11	10.67	9.52	3.63	4.12
	(3.36)	(4.18)	(5.15)	(5.19)	(5.25)	(5.23)	(3.16)	(4.01)
Male household head	0.84	0.78	0.81	0.77	0.82	0.79	0.96	0.87
Persons per bedroom	3.04	3.28	2.42	2.59	2.21	2.45	3.74	3.73
	(1.78)	(2.05)	(1.49)	(1.66)	(1.32)	(1.52)	(1.98)	(2.29)
Connection to public water network	0.31	0.52	0.94	0.89	0.90	0.83	0.08	0.21
Garbage collection by truck	0.25	0.19	0.80	0.81	0.89	0.88	0.05	0.12
Flush toilet	0.37	0.24	0.69	0.68	0.79	0.68	0.27	0.33
Telephone connection	0.04	0.07	0.31	0.27	0.43	0.36	0.00	0.03

	Urban Costa		*Guayaquil*		*Rural Oriente*		*Urban Oriente*	
Indicator	*LSMS*	*Census*	*LSMS*	*Census*	*LSMS*	*Census*	*LSMS*	*Census*
Years of schooling of household head	6.64	6.91	8.88	8.65	5.82	5.16	8.27	8.29
	(4.42)	(4.98)	(4.89)	(4.96)	(3.94)	(4.08)	(4.50)	(4.79)
Male household head	0.83	0.80	0.77	0.78	0.89	0.85	0.83	0.78
Persons per bedroom	3.16	3.12	3.01	2.99	3.54	3.49	2.54	2.64
	(2.08)	(1.96)	(1.87)	(1.92)	(1.89)	(2.26)	(1.50)	(1.63)
Connection to public water network	0.55	0.71	0.72	0.62	0.16	0.29	0.92	0.87
Garbage collection by truck	0.57	0.56	0.75	0.54	0.10	0.20	0.81	0.83
Flush toilet	0.66	0.71	0.76	0.75	0.23	0.18	0.66	0.60
Telephone connection	0.12	0.13	0.25	0.23	0.01	0.03	0.24	0.14

Note: Figures given are means. Standard deviations are in parentheses. All variables except for years of schooling and persons per bedroom are dummy variables taking the value of 1 for a positive response and 0 otherwise.

Source: Authors' calculations.

REFERENCES

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Alesina, Alberto, and Dani Rodrik. 1994. "Distributive Politics and Economic Growth." *Quarterly Journal of Economics* 109(2):465–90.

Angrist, Joshua D., and Alan B. Krueger. 1992. "The Effect of Age of School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples." *Journal of the American Statistical Association* 87(418):328–36.

Arellano, Manuel, and Costas Meghir. 1992. "Female Labour Supply and On the Job Search: An Empirical Model Estimated Using Complementary Data Sets." *Review of Economic Studies* 59(3):537–59.

Atkinson, Anthony. 1989. *Poverty and Social Security*. Hemel Hempstead, U.K.: Harvester Wheatsheaf.

Bramley, Glen, and Gavin Smart. 1996. "Modelling Local Income Distributions in Britain." *Regional Studies* 30(3):239–55.

Bruno, Michael, Martin Ravallion, and Lyn Squire. 1998. "Equity and Growth in Developing Countries: Old and New Perspectives on the Policy Issues." In Vito Tanzi and Ke-Young Chu, eds., *Income Distribution and High-Quality Growth*. Cambridge, Mass.: MIT Press.

Chaudhuri, Shubham, and Martin Ravallion. 1994. "How Well Do Static Welfare Indicators Identify the Chronically Poor?" *Journal of Public Economics* 53(3):367–94.

Deininger, Klaus, and Lyn Squire. 1996. "A New Data Set Measuring Income Inequality." *The World Bank Economic Review* 10(3):565–91.

Elbers, Chris, Jean Olson Lanjouw, and Peter Lanjouw. 2000. "Welfare in Villages and Towns: Micro-Measurement of Poverty and Inequality." Tinbergen Institute Working Paper. (forthcoming).

El Salvador, Government of. 1995. "Priorización de municipios a partir de datos censales." Dirección General de Política Económica y Social, Ministerio de Coordinación del Desarollo Económico y Social. San Salvador.

Farrell, Patrick, Brenda MacGibbon, and Thomas J. Tomberlin. 1997. "Empirical Bayes Estimation Using Logistic Regression Models and Summary Statistics." *Journal of Business and Economic Statistics* 15(1):101–08.

Fields, Gary. 1989. "Changes in Poverty and Inequality in Developing Countries." *The World Bank Research Observer* 4(2):167–85.

FONCODES. 1995. "Mapa de pobreza en Perú." Lima. Processed.

Haddad, Lawrence, and Ravi Kanbur. 1990. "How Serious Is the Neglect of Intra-Household Inequality?" *Economic Journal* 100(402):866–81.

Hentschel, Jesko, and Peter Lanjouw. 1996. "Constructing an Indicator of Consumption for the Analysis of Poverty: Principles and Illustrations with Reference to Ecuador." LSMS Working Paper 124. Development Research Group, World Bank, Washington, D.C. Processed.

Imbens, Guido, and Judith Hellerstein. 1999. "Imposing Moment Restrictions from Auxiliary Data by Weighting." Review of Economics and Statistics 81(1):1–4.

INEI (Instituto Nacional de Estadística e Investigaciones). 1996. "Metodología para determinar el ingreso y la proporción de hogares pobres." Lima. Processed.

Isaki, C. T. 1990. "Small-Area Estimation of Economic Statistics." *Journal of Business and Economic Statistics* 8(4):435–41.

Lanjouw, Jean Olson, and Peter Lanjouw. 1997. "Poverty Comparisons with Non-compatible Data: Theory and Illustrations." Policy Research Working Paper 1709. Policy Research Department, World Bank, Washington, D.C. Processed.

Lusardi, Annemaria. 1996. "Permanent Income, Current Income, and Consumption: Evidence from Two Panel Data Sets." *Journal of Business and Economic Statistics* 14(1):81–90.

Paulin, G. D., and D. L. Ferraro. 1994. "Imputing Income in the Consumer Expenditure Survey." *Monthly Labor Review* 117(12):23–31.

Purcell, N. J., and Leslie Kish. 1980. "Postcensal Estimates for Local Areas (or Domains)." *International Statistical Review* 48(1):3–18.

Ravallion, Martin. 1988. "Expected Poverty under Risk-Induced Welfare Variability." *Economic Journal* 98(393):1171–82.

———. 1994. *Poverty Comparisons*. Chur, Switzerland: Harwood Academic Publishers.

SECAP (Servicio Ecuatoriana de Capacitación). 1994. *Encuesta sobre las condiciones de vida*. Quito.

Sen, Amartya K. 1984. *Resources, Values, and Development*. Oxford: Blackwell.

van de Walle, Dominique, and Kimberly Nead, eds. 1995. *Public Spending and the Poor: Theory and Evidence*. Baltimore, Md.: Johns Hopkins University Press.

World Bank. 1996. "Ecuador Poverty Report." World Bank Country Study. Ecuador Country Department, World Bank, Washington, D.C. Processed.

THE WORLD BANK ECONOMIC REVIEW, VOL. 14, NO. 1: 167-93

Community Targeting for Poverty Reduction in Burkina Faso

David Bigman, Stefan Dercon, Dominique Guillaume, and Michel Lambotte

This article develops a method for targeting antipoverty programs and public projects to poor communities in rural and urban areas. The method calls for constructing an extensive data set from a large number of sources and then integrating the entire set into a geographic information system. The data set includes demographic data from the population census; household-level data from a variety of surveys; community-level data on local road infrastructure, public facilities, water points, and so on; and department-level data on agroclimatic conditions. An econometric model that estimates the impact of household-, community-, and department-level variables on household consumption is used to identify the key explanatory variables that determine the standard of living in rural and urban areas. This model is then applied to predict poverty indicators for 3,871 rural and urban communities in Burkina Faso and to map the spatial distribution of poverty in the country. A simulation analysis assesses the effectiveness of village-level targeting based on these predictions. The results show that such targeting is an improvement over regional targeting in that it reduces leakage and undercoverage.

Budgetary and social pressures to improve the impact of health, education, and rural development programs on the poor are a strong impetus to improve targeting. Undifferentiated transfers that cover the entire population, such as general food subsidies, have proved to be beyond the budget constraints of most devel-

David Bigman is with the International Service for National Agricultural Research, the Netherlands; Stefan Dercon is with the Department of Economics at the Catholic University of Leuven, Belgium, and the Centre for the Study of African Economies at Oxford University; Dominique Guillaume is with the African Department at the International Monetary Fund; and Michel Lambotte is with I-Mage, a consulting group based in Belgium. David Bigman was with the Africa Region at the World Bank while this work was done. The authors' e-mail addresses are *d.bigman@cgiar.org, stefan.dercon@econ.kuleuven.ac.be, dguillaume@imf.org,* and *info@i-mage.be.* This study is part of a project on geographical targeting in Burkina Faso. The authors would like to thank the Belgian Trust Fund for providing financial support and the following institutions in Burkina Faso for providing assistance and data: the Institut National de Statistique et de Démographie (INSD), Famine Early Warning System (FEWS), and many other government departments and institutions including Institut Géographique du Burkina Faso, Direction Générale de l'Hydraulique, Direction des Statistiques Agro-pastorales, Office National de l'Elevage et de l'Agriculture, Direction National de la Météorologie, Ministère de l'Enseignement de Base et de l'Alphabétisation, Programme National de Gestion de l'Information sur le Milieu, Programme National de Gestion des Terroirs, and Direction des Marches Publics. The authors are grateful to Bart Capéau, Dirk Van de Gaer, Jesko Hentschel, Peter Lanjouw, Pramila Krishnan, three anonymous referees, and participants in the World Bank workshop Geographical Targeting for Poverty Reduction and Rural Development for useful comments.

oping countries. Further, the benefits of such transfers go disproportionately to the nonpoor.[1] In many developing countries, particularly those in Sub-Saharan Africa, targeting criteria that cover large geographic areas or large population groups also are likely to be ineffective and too costly to implement. A program that is targeted to the entire rural population, for example, may cover the majority of the country's population—both poor and nonpoor.

More accurate targeting requires criteria that can effectively identify eligible recipients. Such criteria can be narrowly defined, for example, at the level of individual households. Or they can be more broad-based, at the level of a region or province, by identifying the geographic areas or population groups that have a higher than average incidence of poverty (van de Walle 1995). Targeting at the household level is very information-intensive and thus very costly. Identifying eligible households requires complex and expensive means testing. But means-testing is only partly successful even in many industrial countries—despite the wide range of household data that are available in those countries—and a large portion of the benefits leak to ineligible households.

Most developing countries do not have the needed information on individual households, particularly poor households, which limits the scope for narrow targeting. As an alternative to direct means testing, standard household income and expenditure surveys, such as the World Bank's Living Standards Measurement Study (LSMS) surveys, can be used to identify more general characteristics of the poor and thereby to determine a set of indicators, such as number of children or place of residence, that can distinguish the poor.

With the LSMS data set for Côte d'Ivoire, Glewwe (1991) examines the trade-off between using a refined and exhaustive set of indicators for narrow targeting and the cost of collecting the information needed to derive these indicators. He concludes that, for Côte d'Ivoire, a fairly limited set of community and household indicators proved to be quite effective in identifying poor households. However, the incentives for households to change or to lie about their characteristics in order to qualify for a program once these indicators are set can significantly reduce its effectiveness and blow up budgetary costs.[2] This problem, together with the high cost of administering a program at the household level and the danger that these eligibility criteria will leave out many of the country's poor, has deterred the governments of most developing countries from targeting social welfare programs to individual households.

Geographical targeting at the level of the province or region may be an effective approach for reaching the poor in countries where there are substantial disparities in living conditions between geographic areas and where administering poverty programs is relatively straightforward because the local administration is already in place. In India the central government disburses funds across states, in part according to the large disparities in poverty among states. The decision to

1. For simulated examples from Latin America see Baker and Grosh (1994).
2. See also Besley and Kanbur (1991).

locate rural development projects in poorer regions has become the center of India's poverty-oriented agricultural development strategy.

However, even in countries where the poor are concentrated in certain states or regions, geographical targeting at the level of large administrative areas is likely to entail considerable leakage to the nonpoor who live in target areas and fail to cover the poor who live in other areas. Although programs targeted at high levels of geographic aggregation are likely to result in less leakage and better coverage than general, nontargeted programs, their effectiveness in terms of poverty reduction tends to be quantitatively small (Ravallion 1993, Baker and Grosh 1994, Ravallion 1996). Targeted programs may also give households an incentive to move to targeted areas, defeating the purpose of the program and raising its costs.

Ravallion (1993) evaluates the costs and effects of geographical targeting at the level of the province in Indonesia. He concludes that, although targeting clearly helps to alleviate poverty, the magnitude of the gain is small. Baker and Grosh (1994) analyze geographical targeting in Republica Bolivariana de Venezuela, Mexico, and Jamaica and conclude that targeting priority regions can be an effective mechanism for transferring benefits to the poor. But with a given budget constraint, poverty reduction is greater the narrower are the target areas. The greatest reduction in poverty is achieved when the target areas are municipalities or villages.

Narrow geographical targeting at the level of the village or the urban community can reduce leakage in countries or regions where the majority of the population in a village or urban community faces similar socioeconomic conditions and living standards. Many of the households in a village or urban community may have similar sources of income, and all households are affected by the same agroclimatic and geographic conditions, including road conditions, distance to the nearest town, and availability of public facilities for services such as health, education, and water supply. Consequently, income inequality often is due, to a considerable degree, to income differences among villages and only to a lesser degree to income differences among individuals within villages.

But targeting at the district or village level requires much more information on the spatial distribution of poverty across districts or villages and on the characteristics of the poor in these areas. In most developing countries information on the standard of living is provided by a household survey. But the sample size of the standard survey is far too small to allow an estimation of the incidence of poverty at the level of the village or the district for the entire country. At present, the LSMS surveys can provide a map of the spatial distribution of poverty only among the country's main regions.

Some countries that rely on geographical targeting establish criteria for targeting that instead are based on more readily available indicators, such as access to public services, the percentage of school-age children who attend school, and the prevalence of certain illnesses associated with malnutrition. All too often, however, these indicators are not sufficiently correlated with the welfare indicators of

the local population, which may lead to targeting errors in determining eligibility (Hentschel and others 1998).

The objective of this article is to present a method for geographical targeting at the level of rural villages and urban communities. We construct a large data set from several sources and integrate the data at the level of the village or urban community using geo-referencing. We then organize this database in the form of a geographic information system (GIS). With the GIS database we can generate a mapping of poverty at the level of the community and the province. The database includes several strata of information: demographic and socioeconomic information at the household level (taken from a variety of surveys); village- and community-level information, including demographic information from the population census, distance to urban centers, condition of road infrastructure, availability and quality of public services, and sources of drinking water; and departmental or regional information on agroclimatic and geographic conditions, including the location of the main towns and main transport routes.

In the second step of the analysis we use the GIS database, along with detailed data from a household survey, to construct a prediction model of household welfare. The model includes household-, community-, and department-level variables selected from the GIS database and therefore includes only variables for which mean values are available for all communities in the country.

In the third step we use the predictions of this model to estimate the incidence of poverty and the average level of well-being of the households in a community for all communities in the country. These estimates determine, in turn, the spatial distribution of poverty in the country at the village level. We apply this method to Burkina Faso, using the relatively detailed household data of the Priority Survey.

I. The Data

We collected data for this study from a large number of sources and aggregated them at the level of the village according to the name of the village and its geographic coordinates. Some of the data, most importantly the census data, cover all villages in the country or the entire population; other data cover only a sample of villages and a sample of households within each village (table 1). We could not use all of the data in the econometric analysis, however; some of the data did not cover all provinces, while other data, most notably those from the agricultural survey, did not contain the information needed to incorporate them into the GIS database.

After we collected the data, we standardized them and integrated them within a common data set. This data set contained more than 60 tables with information on the geographic coordinates of all villages, towns, markets, and public facilities; data on the entire road network; socioeconomic and demographic data from a variety of surveys and the population census; and data on the agroclimatic conditions in the country's main provinces. We then organized the data as a GIS, that is, a computer system that allows the analysis and display of geographic and other data.

Table 1. *Data Sources*

Level of aggregation	*Source*	*Coverage*
Household	Priority Survey (1994): data on income and expenditures for 8,642 households	Survey sample (473 villages)
Village	Priority Survey (1994): community component, which covers infrastructure and communal services	Survey sample (473 villages)
Village	National census (1985): demographic data	National
Village	Ministry of Water Management and Infrastructure (1995): data on health and water infrastructure, distances to infrastructure, public administration, and social groupings	25 of 30 provinces
Village	Ministry of Education (1995): data on primary school infrastructure and teacher-pupil ratios	National
Department	Ministry of Agriculture (1993): data on various indicators ranging from average literacy rates to vegetation indexes	National
Department	Directorate of Meteorology (1961–95): data on temperature (31 locations), evapotranspiration (15 locations) and rainfall (160 locations)	National
Province	Ministry of Agriculture (1993): data on cattle per household	National

To illustrate the type of information that we extracted from the GIS for the community study, figure 1 (in the appendix) shows the location of water points and their proximity to the villages in the Department of Karangasso-Vigue. The points on the map that indicate the location of the villages are scaled according to the size of the village population, thus showing the demand pressures on each water point. The map also contains information on road infrastructure, including the quality of roads, and on the hydrographic networks.

We had to limit our study to a smaller data set, because not all data were available for all villages at the time that the data were collected for inclusion in the GIS. In particular, the data obtained from the Ministry of Water Management were limited to 25 provinces, or 5,207 of the country's 6,821 villages. Data for the remaining five provinces were subsequently collected from the Ministry in another survey, but very few variables were comparable between the two surveys. In some villages data on other variables were also missing, and thus we had to reduce the number of villages in the final prediction analysis to 3,871, or 57 percent of the country's total number of villages (descriptive statistics are given in tables 2 and 3).

The lack of sufficient data for all 6,821 villages is, of course, a cause for concern. Some of the missing data are available in the archives of the different ministries and, in principle, could be retrieved. However, if a significant number of villages still do not have all the necessary data, targeting will have to be made at higher levels of geographic aggregation (at the department or province level). To target at these higher levels, we still would have to use the predictions of village-level poverty obtained for all villages outside the sample; we cannot use the household survey directly because the sampling frame and the sample size do not adequately represent all departments and provinces.

Table 2. *Descriptive Statistics on Variables Used in the Estimation*

Aggregation level	Variable	Urban Mean[a]	Urban Standard error[a]	Urban Number of observations	Rural Mean[a]	Rural Standard error[a]	Rural Number of observations	Data source
Household	Children 0–6 years per adult (15–50 years) in household	0.530	0.495	2,671	0.779	0.598	5,508	Priority Survey
Household	Children 7–14 years per adult in household	0.618	0.590	2,671	0.748	0.640	5,508	Priority Survey
Household	Elderly persons (50+) per adult in household	0.183	0.343	2,671	0.313	0.426	5,508	Priority Survey
Household	Literate head in household	0.477	0.499	2,736	0.134	0.341	5,906	Priority Survey
Household	Percentage male adults literate in household	0.562	0.422	2,736	0.177	0.313	5,906	Priority Survey
Household	Percentage female adults literate in household	0.373	0.397	2,736	0.053	0.174	5,906	Priority Survey
Household	Livestock units per capita	0.123	0.909	2,736	0.442	0.943	5,906	Priority Survey
Village	Distance to nearest rural primary school	n.a.	n.a.	n.a.	2.290	5.640	4,412	Ministry of Water Management and Infrastructure
Village	Teachers per child ages 7–14 years	0.014	0.002	2,736	0.005	0.006	5,760	Ministry of Education
Village	Distance to nearest health facility	n.a.	n.a.	n.a.	4.790	7.770	4,434	Ministry of Water Management and Infrastructure

Village	Nearest facility has safe water	0.820	0.390	2,416	0.034	0.180	4,434	Ministry of Water Management and Infrastructure
Village	Number of pumps per rural community	n.a.	n.a.	n.a.	7.350	10.640	5,241	Ministry of Water Management and Infrastructure
Village	Presence of an all-weather road	n.a.	n.a.	n.a.	0.570	0.500	4,434	Ministry of Water Management and Infrastructure
Department	Cultivated area in department per capita	0.211	0.221	2,736	0.507	0.301	5,760	Famine Early Warning System
Department	Average rainfall, 1980–94	65.800	10.070	2,736	62.500	14.840	5,760	Directorate of Meteorology
Department	Absolute value of deviation of rainfall from average, 1994	19.450	14.490	2,736	22.580	12.960	5,760	Directorate of Meteorology
Department	Average length rainy season, 1982–92	9.520	1.340	2,736	9.530	2.000	5,760	Famine Early Warning System
Department	Average vegetation index, 1982–92	0.114	0.034	2,736	0.136	0.051	5,760	Famine Early Warning System
Department	Homogeneity of rainy season, 1982–92	0.162	0.019	2,736	0.161	0.036	5,760	Famine Early Warning System

n.a. Not applicable.
Note: For village- and department-level variables, the same value is assumed for all households in the community.
a. Weighted using sampling weights.
Source: Authors' calculations.

Table 3. *Descriptive Statistics on Variables Used in the Prediction*

		Urban			*Rural*			
Aggregation level	*Variable*	*Mean*[a]	*Standard error*[a]	*Number of observations*	*Mean*[a]	*Standard error*[a]	*Number of observations*	*Data source*
Village	Children ages 0–6 years per adult (15–50 years) in household	0.656	0.110	300	0.645	0.227	6,818	National Census
Village	Children ages 7–14 years per adult in household	0.593	0.120	300	0.563	0.280	6,818	National Census
Village	Elderly persons (50+) per adult in household	0.320	0.076	300	0.348	0.351	6,818	National Census
Province	Literate head in household	0.450	0.181	191	0.113	0.075	6,711	Priority Survey
Province	Percentage male adults literate in household	0.522	0.147	191	0.141	0.079	6,711	Priority Survey
Province	Percentage female adults literate in household	0.323	0.149	191	0.044	0.034	6,711	Priority Survey
Province	Livestock units per capita	0.147	0.090	191	0.492	0.263	6,711	Ministry of Agriculture
Village	Distance to nearest rural primary school	n.a.	n.a.	n.a.	4.390	5.040	4,556	Ministry of Water Management and Infrastructure
Village	Teachers per child ages 7–14 years	0.023	0.032	295	0.003	0.011	4,753	Ministry of Education
Village	Distance to nearest health facility	n.a.	n.a.	n.a.	6.790	7.460	4,393	Ministry of Water Management and Infrastructure

Village	Nearest facility has safe water	0.150	0.350	219	0.005	0.073	4,390	Ministry of Water Management and Infrastructure
Village	Number of pumps per rural community	n.a.	n.a.	n.a.	2.350	2.765	5,425	Ministry of Water Management and Infrastructure
Village	Presence of an all-weather road	n.a.	n.a.	n.a.	0.430	0.500	4,618	Ministry of Water Management and Infrastructure
Department	Cultivated area in department per capita	0.669	0.605	300	0.751	0.717	6,821	Famine Early Warning System
Department	Average rainfall, 1980–94	65.520	15.340	300	69.160	16.340	6,821	Directorate of Meteorology
Department	Absolute value of deviation of rainfall from average, 1994	18.900	11.540	300	18.610	13.950	6,520	Directorate of Meteorology
Department	Average length rainy season, 1982–92	10.190	2.310	300	10.770	2.440	6,520	Famine Early Warning System
Department	Average vegetation index, 1982–92	0.126	0.054	300	0.121	0.051	6,821	Famine Early Warning System
Department	Homogeneity of rainy season, 1982–92	0.152	0.038	300	0.153	0.036	6,821	Famine Early Warning System

n.a. Not applicable.
Note: For department-, province-, and village-level variables, the same value is assumed for all households in the community.
a. Weighted using total population relative to village population.
Source: Authors' calculations.

Similar to the LSMS surveys, the sampling for the Priority Survey used to estimate the consumption model was semistratified (INSD 1996). The survey was designed to be representative at both the national and regional levels. The country was divided into seven regions: five rural regions representing five agroclimatic areas and two urban regions, one comprising Ouagadougou and Bobo-Dioulasso, Burkina Faso's two main cities, and the other comprising the remaining cities. From the seven regions 434 enumeration areas were selected on the basis of their socioeconomic characteristics. In each, 20 households were randomly selected. For our econometric estimation, however, we had to reduce the sample size to 5,618 households and 201 enumeration areas because of missing data for five provinces and incomplete data for certain variables in a few other villages.

II. METHODOLOGY

The econometric analysis in this article has two parts. The first estimates a prediction model for household consumption, using the household data of the Priority Survey and the community data from all other sources, in order to determine the variables that best explain household consumption and poverty. The explanatory variables we select for the prediction model include only those for which we have data for all villages outside the Priority Survey sample. In the second part of the analysis we use the prediction model and the village-level data from the GIS database to measure welfare at the village level for all villages outside the Priority Survey sample. In line with similar studies on this subject, we use consumption per standard adult (adult equivalent) as our welfare indicator at the household level and use the headcount index as the measure of poverty.[3]

The Two Models

Let c_{ij} denote the level of consumption per standard adult in household i located in community j. Let z denote the poverty line, and let $y_{ij} = c_{ij}/z$ be the normalized welfare indicator per standard adult. We conduct the analysis in terms of the natural logarithm of y_{ij}. For a poor person, therefore, $y_{ij} < 1$, or $\ln y_{ij} < 0$. The headcount index, H_j, which measures the relative size of the poor population in community j, is equal to the mean value of the individual poverty indicators, H_{ij}, which indicate the probability that the household ij is poor, over all the individuals in that community. The individual poverty indicator is determined by the normalized welfare function as follows:

$$H_{ij} = 1 \text{ if } \ln y_{ij} < 0 \qquad (1)$$
$$H_{ij} = 0 \text{ if } \ln y_{ij} \geq 0.$$

In constructing the prediction model, we represent the individual welfare indicator as a function of a vector of household and community explanatory variables $\mathbf{X}_{ij}$ and a residual term u_{ij}, which is assumed to be normally distributed with

3. We use simple nutritional adult equivalence scales. These are 0.3 for a child less than 5 years old and 0.7 for a child between 5 and 15 years. Each adult counts as 1.0.

$u_{ij} \sim N(0,\sigma_j^2)$, thereby allowing for village-level heteroskedasticity. The prediction model is thus given by:

$$\ln y_{ij} = \boldsymbol{\beta}' \mathbf{X}_{ij} + u_{ij}. \tag{2}$$

As noted earlier, we only select explanatory variables if their mean values are available for all villages in the GIS database. They include community characteristics and mean household characteristics, such as household composition and literacy rates, for all households in the community. We can estimate equation 2 using the maximum likelihood method, with $u_{ij} \sim N[0, \sigma^2\exp(\gamma'\mathbf{X}_j^V)]$, where $\mathbf{X}_j^V$ are the mean values of the explanatory variables in community j. This formulation corrects for heteroskedasticity and generates the estimators $\mathbf{b}$ and s_j of the parameters $\boldsymbol{\beta}$ and σ_j.

We use these estimators and the set of explanatory variables to predict a community's mean consumption for all communities outside the Priority Survey sample. Mean consumption, however, is not necessarily a good predictor of poverty, since the poverty measure is a function of not only mean consumption, but also the distribution of consumption within the community. The term s_j represents one part of that distribution, since the within-community variance is the sum of the variance of the regression and the deviation of predicted household consumption from predicted mean consumption.[4]

Using these estimators and the set of explanatory variables, a consistent estimate of the probability that household ij with characteristics $\mathbf{X}_{ij}$ is poor can then be expressed as:

$$E(H_{ij} \mid \mathbf{X}_{ij}, \mathbf{b}, s_j) = \text{Prob}\,(u_i < -\mathbf{b}'\mathbf{X}_{ij}) = \Phi\,(-\mathbf{b}'\mathbf{X}_{ij}/s_j) \tag{3}$$

where $\Phi(.)$ is the cumulative normal distribution. The predicted incidence of poverty in community j is determined from equation 3 as:

$$E(H_j \mid \mathbf{X}_{ij}, \mathbf{b}, s_j) = E\,[\Phi\,(-\mathbf{b}'\mathbf{X}_{ij}/s_j)]. \tag{4}$$

If complete information on $\mathbf{X}_{ij}$ were available for all households and all villages in the country, this prediction would be fairly straightforward: we would use equation 3 to estimate the probability that each household in the village is poor and equation 4 to predict the incidence of poverty in the community—across all villages outside the sample.[5] However, in Burkina Faso the only data available for all villages outside the Priority Survey sample are the *mean* values of the explanatory variables $\mathbf{X}_j^V$ in each community. Since equation 4 is nonlinear, we cannot use these variables simply to predict village-level poverty. But we can use Taylor expansions to obtain an approximation. Thus we expand

4. The within-village variance of consumption can be written as $E\{[Y_{ij} - E(Y_j)]^2\} = E[(\mathbf{b}'\mathbf{X}_{ij} - \mathbf{b}'\mathbf{X}_j^V)^2] + s_j^2$, in which Y_j is the mean level of consumption in the village. In words, the variance of consumption is the sum of the squared deviation of predicted household consumption from predicted mean consumption in each village and the village-level variance of the prediction model.

5. Hentschel and others (1998) use this property to predict regional poverty from census data.

equation 4 around $-\mathbf{b}'\mathbf{X}_j^V/s_j$. Using the property that $E(\mathbf{b}'\mathbf{X}_{ij} - \mathbf{b}'\mathbf{X}_j^V) = 0$, we obtain:

$$E(H_j) = E[\Phi(-\mathbf{b}'\mathbf{X}_{ij}/s_j)] \tag{5}$$

$$\approx \Phi(-\mathbf{b}'\mathbf{X}_j^V/s_j) + \tfrac{1}{2}(\mathbf{b}'\mathbf{X}_j^V/s_j^3)\phi(-\mathbf{b}'\mathbf{X}_j^V/s_j)\, E(\mathbf{b}'\mathbf{X}_{ij} - \mathbf{b}'\mathbf{X}_j^V)^2$$

where $\phi(.)$ is the normal density function and $E(\mathbf{b}'\mathbf{X}_{ij} - \mathbf{b}'\mathbf{X}_j^V)^2$ is the variance of predicted household consumption around predicted mean consumption within each village (Maddala 1983). In words, the predicted level of poverty for villages outside the sample is a function of the mean level of consumption per adult and the variance around that mean. Equation 5 can therefore predict the incidence of poverty in communities outside the sample, using the estimated parameters of the household consumption function ($\mathbf{b}$ and s_j) and the community-level characteristics $\mathbf{X}_j^V$ of the villages outside the sample.

Empirical Estimation

We use the regression analysis to predict consumption levels for all households rather than to determine whether or not a household is poor. The latter approach would be equivalent to estimating equation 1 directly. That approach is often referred to as a multivariate poverty profile (Ravallion 1996). The individual poverty indicator in equation 1 is binary, so one could use a probit (or an alternative) model to construct a prediction model. As Ravallion points out, a puzzling feature of this approach is that the estimation techniques used typically were developed for situations in which the observed data were dichotomous or truncated at zero, even though consumption was observed.

The standard way of solving this estimation problem is to define a regression model in which a continuous latent (unobserved) variable is regressed on a set of observed explanatory variables (Maddala 1983). A particular error structure (such as the normal distribution for the probit) is then assumed, allowing the parameters of interest to be estimated. These parameters can then be used for inference related to the explanatory variables and the observed limited dependent variable. If this procedure is used on a poverty indicator, such as the headcount index, then the latent variable is an observed variable that was used to calculate the limited dependent variable. Since the latent variable is observed, limited dependent variable estimation of the poverty indicator is not necessary and would be less efficient, since some information actually available (consumption) is not used in constructing the prediction model.

To estimate household consumption, we use a standard reduced-form framework in which income (measured in terms of household consumption) is regressed on household characteristics, including human and physical capital, as well as on community characteristics.[6] Some community characteristics are specified at the village level, whereas others, primarily the agroclimatic conditions, are specified at the department or regional level.

6. Examples of this approach are Glewwe and Kanaan (1989) and Coulombe and McKay (1996). Glewwe (1991) has a useful discussion on the justification for including particular variables in this type of approach. We return to the problems related to this specification below.

The Priority Survey contains a limited but important set of variables that can be used to explain household consumption. We select household-level explanatory variables that allow aggregation at the community level and thus can be used for the prediction model. This limited the choice of household-level variables to those for which the corresponding mean values at the community level are available for all communities in the country. As a consequence, we could not include in the estimation several variables, such as the education of household members (as opposed to the literacy of the household head), that usually are found to be significant in a consumption model. Furthermore, data on household assets and landholdings, which are also significant explanatory variables in most consumption models, are not available in the Burkina Faso Priority Survey. This reduced the explanatory power of the model and very likely created a missing variable problem. Moreover, since the underlying data are cross-sectional, household heterogeneity—a common problem affecting any regression of welfare indicators—is also difficult to address. Despite these reservations, we were able to collect data on most of the important explanatory variables, such as demographic variables and variables representing human and physical capital, and include them in the model (see table 2).

We also include village- and department-level variables. Department-level variables are primarily climatic data and means of certain household variables (such as the average area of cultivated land across all households in the department), which we obtained from the Ministry of Agriculture. We distinguish between the impact of long-term climatic characteristics and the impact of temporary fluctuations by including among the explanatory variables the average level of rainfall in the past 15 years and the absolute value of the deviation of the previous year's rainfall from the long-term average. The village-level explanatory variables also include data on the distance to and quality of schools and health facilities, the quality of access roads, and the quality of the water supply.

III. Estimating Poverty within the Sample

Descriptive statistics on poverty and consumption reveal large differences in the incidence of poverty between urban and rural households and a much higher standard of living in the country's two main cities (table 4). Among rural areas, the West has the lowest rates of poverty. Among the other regions, differences between poverty rates are small.

In the econometric analysis we regress consumption per adult equivalent on the explanatory variables (listed in table 2), according to the linear model given by equation 2. We estimate this model using the maximum likelihood method, in which the regression coefficients and the heteroskedastic errors are estimated jointly (tables 5 and 6).[7] In allowing for heteroskedasticity by community, we can use the community-level information to predict mean consumption per equiva-

7. The regression was weighted with individual sampling weights derived from the original sampling frame used by the World Bank and INSD.

Table 4. *Poverty and Consumption in Burkina Faso*

Region	*Consumption per adult equivalent (Francs CFA per month)*	*Headcount index*[a]
Rural		
West	7,573	0.56
South/South-East	5,699	0.67
Central-North	4,952	0.74
Central-South	5,240	0.75
North	6,122	0.64
Bobo-Dioulasso and Ouagadougou	20,768	0.14
Other urban	12,173	0.39
National	8,766	0.58

a. The poverty line is set at two-thirds of mean consumption.
Source: Priority Survey (1994).

lent adult as well as the variation around this mean. The estimated variance within a community may provide some information on the extent of inequality in the distribution of consumption. We conduct the regression analysis separately for households in rural and urban areas (pooling tests convincingly reject running one national regression). For both urban and rural areas multiplicative heteroskedasticity cannot be rejected at the 1 percent level (table 6).[8]

Because we use maximum likelihood estimation to jointly determine the coefficients in the model and the structure of heteroskedasticity, no simple R^2 can be reported. However, the ordinary least squares (OLS) estimates of the model (derived in the first step) indicate that the adjusted R^2s are low, equal to 0.28 for the urban population and 0.17 for the rural population. These figures are low primarily because of restrictions on the choice of variables included in the model. When we use all the household- and community-level variables that were available in the Priority Survey in the estimation, the R^2 in the regression for rural households rises to 0.50. Because of the low values of the adjusted R^2s, we need to make significant adjustments in applying the results in the prediction model. These adjustments are discussed in section IV.

Interpreting the Results

The results show that the variables included in the model are strongly jointly significant and that a substantial number of household- and community-level variables are highly significant. The following results stand out.

The household variables that are correlated most closely with the level of consumption in both rural and urban areas are the adult literacy rates. (The variables describing adult literacy also include literacy of the household head.) The

8. The Breusch-Pagan Lagrange-multiplier test convincingly rejects homoskedasticity (see table 6). The Glesjer (1965) test indicates that in both urban and rural areas the null hypothesis of multiplicative heteroskedasticity cannot be rejected at the 1 percent level.

Table 5. *Regression Results: Poverty and Consumption*

	Rural		*Urban*	
Variable	*Coefficient*	*t-value*	*Coefficient*	*t-value*
Constant	7.71	52.48*	10.82	21.99*
Children 0–6 years per adult (15–50 years) in household	0.02	1.55	0.01	0.40
Children 7–14 years per adult in household	–0.03	–1.67***	–0.04	–1.60
Elderly persons (50+ years) per adult in household	0.03	1.24	–0.13	–2.91*
Literate head in household	0.18	3.66*	0.33	7.63*
Percentage male adults literate in household	0.13	2.48**	0.16	3.18*
Percentage female adults literate in household	0.55	8.41*	0.42	10.11*
Livestock units per capita(/10)	0.93	11.06*	0.31	1.49
Distance to nearest rural primary school (/100)	–0.48	–2.80*		
Teachers per child ages 7–14 years (*10)	0.21	1.07	–0.39	–0.22
Distance to nearest health facility (*100)	0.18	1.58		
Nearest facility has safe water	0.14	2.92*	0.92	5.24*
Number of pumps per rural community (/100)	0.34	3.27*		
Presence of all-weather road	0.10	4.83*		
Cultivated area in department per capita	0.01	0.32	1.66	5.29*
Average rainfall, 1980–94 (/100)	0.53	3.39*	–1.64	–2.28**
Absolute value of deviation of rainfall from average, 1994 (/100)	–0.22	–2.78*	–0.44	–2.03**
Average length of rainy season, 1982–92	–0.01	–0.69	–0.06	–0.98
Average variable vegetation index, 1982–92	–0.54	–1.81***	8.17	3.24*
Homogeneity of rainy season, 1982–92	2.50	8.38*	–12.26	–3.63*
F-joint significance regression	34.58* *F*[19, 4107]		53.70* *F*[15, 2346]	
Number of valid observations	4,119		2,362	

***Significant at the 10 percent level.
**Significant at the 5 percent level.
*Significant at the 1 percent level.
Note: Dependent variable is log of consumption per standard adult.
Source: Authors' calculation.

dependency rates, namely the number of children and elderly persons per adult in the household, do not have a clear effect on consumption in rural households. But in urban households the number of elderly persons per adult in the household has a significant impact on the level of consumption per adult equivalent.

The number of livestock units per capita—the only proxy for the household's physical assets available in the Priority Survey—is significantly and positively correlated with consumption in rural areas. The community-level variables that characterize agroclimatic conditions also have a strong impact on consumption in rural areas. Consumption per capita typically is higher in rural areas that have relatively high levels of long-run average rainfall, relatively normal rain in the survey year, and low rainfall variability over the rainy season.

However, the agroclimatic variables representing the average level of rainfall and the homogeneity of the rainy season seem to have a negative effect on consumption in urban areas. A possible explanation is that the consumption basket

Table 6. *Regression Results: Estimated Variance with Multiplicative Heteroskedasticity*

	Rural		*Urban*	
Variable	*Coefficient*	*t-value*	*Coefficient*	*t-value*
Constant	0.12	5.03*	0.03	1.75***
Children 0–6 years per adult (15–50 years) in household (community mean)	0.40	2.85*	–0.86	–3.68*
Children 7–14 years per adult in household (community mean)	0.79	6.27*	1.12	5.08*
Elderly persons (50+) per adult in household (community mean)	–0.29	–1.64***	0.47	1.53
Literate head in household (community mean)	0.49	3.86*	0.09	0.91
Percentage male adults literate in household (community mean)	–0.26	–1.96**	0.12	1.12
Percentage female adults literate in household (community mean)	0.11	0.90	0.05	0.70
Livestock units per capita (community mean)	–0.05	–0.94	0.00	0.04
Distance to nearest rural primary school	0.00	0.07		
Teachers per child ages 7–14 years (*100)	0.20	4.70*	2.00	4.95*
Distance to nearest health facility	0.00	–0.97		
Nearest facility has safe water	–0.38	–3.28*	–1.47	–3.63*
Number of pumps per rural community	0.01	2.16**		
Presence of all-weather road	0.20	4.03*		
Cultivated area in department per capita	0.33	3.05*	–1.03	–1.33
Average rainfall, 1980–94	0.02	4.20*	0.02	1.38
Absolute value of deviation from average, 1994	0.00	1.14	0.00	–0.38
Average length rainy season, 1982–92	0.09	2.78*	0.00	–0.01
Average variable vegetation index, 1982–92 (*10)	0.31	4.36*	–3.19	–5.56*
Homogeneity of rainy season, 1982–92 (*10)	–0.16	–2.25**	4.35	5.65*
Breusch-Pagan LM heteroskedasticity	603.35** (19 degrees of freedom)		158.33 (15 degrees of freedom)	
Glesjer test multiplicative heteroskedasticity	3.66* *F*[19, 4107]		3.66* F[15, 2328]	

***Significant at 10 percent level.
**Significant at 5 percent level.
*Significant at 1 percent level.
Note: A positive coefficient on the explanatory variable indicates that this variable has the effect of raising the variance; a negative coefficient indicates that this variable has the effect of lowering the variance.
Source: Authors' calculations.

of urban households typically includes commodities that were not recorded in the Priority Survey (which includes only a small number of consumption items); the reduction in consumption during the normal years recorded in the survey is therefore spurious. Another possible explanation is that these variables are correlated with significant missing variables that have a negative impact on the consumption of urban households. The data we had at our disposal did not allow us to analyze these effects further.

In rural areas consumption in villages that are farther away from schools is generally lower. No similar effect is revealed with respect to the distance to health facilities. One explanation may be that in some regions villages located farther from a health facility receive services from mobile health clinics. In both urban and rural areas the quality of services in the health facility—approximated by the availability of safe drinking water in the facility—is significantly correlated with the level of per capita consumption in the surrounding villages and urban neighborhoods. Only about one-third of the health facilities in Burkina Faso have safe drinking water.

In rural areas the quality of infrastructure, indicated by the availability of safe drinking water in the village (measured by the number of functioning pumps), and the quality of access roads to the village have a significant and positive impact on consumption. Mean consumption in villages with an all-weather access road is nearly 10 percent higher than in villages without one. The greater opportunity to trade, rather than produce for own-consumption, and the better alternatives for nonagricultural work that access to an all-weather road provides are the main reasons for this effect.

The coefficients that determine the pattern of the village-level error terms indicate that in villages where a relatively high proportion of household heads are literate, the distribution of per capita consumption is less equal than in villages where a low proportion of household heads are literate. It may be that in a village with more literate adults the income differences between households with less educated heads and households with more educated heads are relatively larger.

Households in villages with relatively high average levels of rainfall differ more widely in per capita consumption, possibly because in these villages some households are better equipped and more able to take advantage of superior agricultural conditions. Villages with higher average landholdings per household show larger variability in per capita consumption.

Accounting for Endogeneity

Several of the interpretations of the results suggest possible causal relationships between the explanatory variables and the dependent variable. However, these interpretations are intended primarily as background for a more thorough evaluation of the possible policy implications. We make them with the usual caveat of potential endogeneity of the community-level variables, which means that correlation need not imply causality. Thus, for example, the government's policy of locating more public education facilities in poorer villages as part of its antipoverty program will lead to a high negative correlation between average per capita consumption and the proximity of the village to a school (Rosenzweig and Wolpin 1986). To take another example, the availability of a large number of water pumps in a village need not be the cause of a relatively high standard of living, but rather may be the result of a higher demand for safe drinking water among more affluent villagers.

The purpose of our regression estimates is, however, to construct a prediction model that can identify the poorest and wealthiest villages. The quality of such predictions depends only on the degree of correlation between the explanatory and the dependent variables, irrespective of whether or not the correlation indicates causality. If, for example, health facilities are intentionally placed in poorer villages, then the distance from a village to a health facility can be useful for predicting the standard of living.[9] Nevertheless, the possibility of endogeneity forces us to use special care in interpreting the results for policy purposes.[10] Although the significance and size of the coefficients are suggestive, additional work is needed to design appropriate policies for reducing poverty.

IV. Predicting the Geographic Distribution of Poverty

In the next step we apply the regression results (obtained for a sample of communities) to the data available in the GIS database in order to predict the distribution of poverty across all communities in Burkina Faso. We focus on the 3,871 villages, out of a total of more than 6,000, for which all the necessary information was available. We calculate the headcount index of a community from equation 5, using the parameter estimates that were obtained in the regression analysis. To use this equation, we also need estimates of mean consumption per adult in the community and the variance of consumption. For the term $E(\mathbf{b}'\mathbf{X}_{ij} - \mathbf{b}'\mathbf{X}_j^V)^2$ in equation 5 we use the average value per region (rather than per community) in the survey data, and we obtain s_j using the coefficients given in table 6. We predict mean consumption per standard adult in the communities outside the sample from the mean values of the explanatory variables for each of these communities, using the coefficients given in table 5.

Before applying these predictions for all communities outside the sample, we assess their quality by comparing them with direct estimates of poverty from the sample of 201 communities included in the Priority Survey. We derive the correlation coefficients of the predicted and calculated level of poverty for these villages. The value of the Pearson-correlation coefficient is 0.51, and it is strongly statistically significant. For policy decisions, however, the more relevant criterion is the order of villages along the poverty scale. To test this aspect of the prediction model, we calculate the Spearman rank correlation coefficient between the order established by the direct estimates of poverty and the order established by the model's predictions. That coefficient is also strongly significant at 0.43.

Another way to test the quality of the predictions is to compare the estimated and the predicted values of the headcount index. We do this for selected communities in three provinces (table 7). Although the predicted values of the headcount index in each community often fall outside the confidence interval of the calcu-

9. This assumption also requires that the same program placement rule be used outside the sample as inside the sample. Since the sample is nationally representative, this may be an appropriate assumption.

10. There are other possible sources of endogeneity. For example, we assume that location is not a choice variable. We therefore do not consider migration explicitly.

Table 7. *Comparison of Predictions and Direct Estimates of the Headcount Measure of Poverty for Sample Villages in Three Provinces*

Province	*Village identification number*	*Within-sample estimates*[a]	*Outside-sample predictions*
Kossi	4426	0.28 (0.01)	0.24
	3786	0.33 (0.01)	0.67
	512	0.54 (0.01)	0.64
	2936	0.54 (0.04)	0.65
	5266	0.57 (0.02)	0.56
	5117	0.64 (0.02)	0.69
	1626	0.68 (0.02)	0.70
	1556	0.69 (0.01)	0.61
	1290	0.78 (0.01)	0.70
	250	0.80 (0.01)	0.57
Kouritenga	744	0.64 (0.03)	0.66
	6233	0.72 (0.03)	0.66
	657	0.75 (0.01)	0.57
	1627	0.80 (0.03)	0.74
	2943	0.83 (0.01)	0.65
	3213	1.00 (0.00)	0.64
	3828	1.00 (0.00)	0.76
Mouhoun	1278	0.23 (0.01)	0.32
	790	0.36 (0.02)	0.52
	6753	0.48 (0.03)	0.57
	4982	0.48 (0.02)	0.56
	740	0.50 (0.02)	0.52
	5149	0.57 (0.02)	0.52
	5635	0.72 (0.01)	0.60
	6674	0.72 (0.01)	0.65

a. Standard errors are in parentheses (from Deaton 1997:47). The figures are weighted by household size.
Source: Authors' calculations.

lated measure, the rank order of communities from richest to poorest in each province is similar.

Despite these results, the low R^2 values in the regression analysis and the poor quality of the data prevent us from using these predictions directly. Moreover, these predictions rely on the assumption of normality of the error term. One common test for normality is the Jarque-Bera test. Our estimate of the Jarque-Bera statistic is 11.8, and we therefore have to reject the normality hypothesis.[11]

As a result, we do not use the prediction to establish a complete order of communities on the poverty scale. Instead, we divide the 3,871 villages and urban communities into four categories of poverty, ranging from the poorest to the wealthiest, according to predicted levels of poverty. Despite the errors in these predictions, our results suggest that most of the villages categorized as "poorest"

11. This statistic has a chi-square distribution with two degrees of freedom. The normality hypothesis has to be rejected at a probability of 0.997. However, the Jarque-Bera test is not robust to the presence of heteroskedasticity, which could not be rejected by the Breusch-Pegan LM test and the Glesjer test. We are not aware of a test of normality in the presence of heteroskedasticity, but the high value of the Jarque-Bera statistic suggests that it is highly probable that the residuals are not distributed normally.

are likely to have a higher incidence of poverty than most of the villages categorized as "least poor." The villages in the poorest category are therefore candidates for targeted poverty alleviation programs, and the villages in the least poor category are candidates for cost-recovery programs.

Given the data limitations in Burkina Faso, effective targeting would have to focus only on these two extreme categories in order to reduce leakage as much as possible and keep within the budget constraint. We can further improve targeting in the present circumstances—given the limited availability and poor quality of the data—by dividing the villages into a larger number of categories and concentrating on villages in the two extreme categories. Future research intended to improve targeting will have to focus, however, on efforts to improve the quality of the data as well as generate additional series of geo-referenced data.

We construct the geographic distribution of rural and urban communities across these categories of well-being within each province (table 8). We divide the villages into the four categories, ranging from poorest to least poor, using the predicted values of poverty incidence and allocating the entire population in each village to the corresponding category. We set the categories so that the popula-

Table 8. *Distribution of Communities by Poverty Category*
(percent)

	Poverty category				*Percentage*
Province	*Poorest*	*Lower-middle*	*Upper-middle*	*Least poor*	*of total population*
Bam	0	10	36	54	0.90
Bazega	13	36	35	16	4.85
Boulgou	37	22	32	9	7.04
Boulkiemde	41	36	15	7	5.99
Ganzourgou	47	33	19	2	2.74
Gnagna	24	13	21	41	3.73
Gourma	34	17	23	25	5.55
Kossi	26	30	29	14	6.22
Kouritenga	7	25	43	25	3.11
Mouhoun	41	39	18	2	6.06
Nahouri	23	29	40	8	1.71
Namentenga	4	25	18	52	2.16
Oubritenga	9	24	40	27	4.73
Oudalan	24	36	19	21	1.69
Passore	11	2	18	68	3.77
Sanguie	31	23	25	21	4.64
Sanmatenga	17	11	25	46	7.64
Seno	26	29	19	26	4.70
Sissili	41	43	16	0	3.69
Soum	3	14	21	61	2.54
Sourou	12	12	12	64	5.31
Tapoa	60	17	22	1	2.12
Yatenga	24	24	31	2	6.72
Zoundweogo	3	59	30	8	2.40

Note: The poverty line is set at two-thirds of mean consumption.
Source: Authors' calculations.

tion in each represents 25 percent of the country's total population. The distribution of the population within provinces is significantly different, however. For example, 41 percent of the population in the province of Boulkiemde lives in villages classified as poorest, and only 7 percent of the population lives in villages classified as least poor.

Consider, as an illustration, an antipoverty program targeted to the five provinces in which at least 40 percent of the population resides in villages classified as poorest. Under this criterion 21 percent of the country's population will be included in the target provinces. Only 3 percent of the population in the five target provinces (which account for only 0.6 percent of the country's total population), however, lives in villages classified as least poor, suggesting that leakage is likely to be small. Nearly 43 percent of the population in the five target provinces lives in villages classified as poorest, accounting for 36 percent of the country's population that lives in the poorest villages. At the other extreme, a cost-recovery program that is targeted to the seven provinces in which more than 40 percent of the population lives in villages classified as least-poor will cover 26 percent of the total population but only 13 percent of the population that lives in the poorest villages.

Targeting antipoverty programs at the province level, however, is likely to be less effective than targeting at the village level. The reason is that targeting provinces is bound to include villages of the higher categories in which the incidence of poverty is likely to be less than that in villages of the lowest category. Therefore, under a given budget constraint, a targeted program at the village level that focuses only on villages categorized as poorest is likely to cover a larger share of the country's poor population and entail less leakage, despite the prediction errors in classifying villages.

Most of the urban communities are classified as least poor, largely because of the much higher standard of living and much lower incidence of poverty in urban areas. There are also several other, more technical, explanations. In urban areas the distinction between poor and nonpoor communities is less clear than in rural areas. In many developing countries it is not uncommon for poor households to reside in relatively affluent urban communities and vice versa. Further, urban communities, as defined in the household surveys, are, in fact, enumeration areas that have been demarcated by local authorities for administrative purposes, and their borders are often quite arbitrary. Whereas in rural areas enumeration areas are generally limited to one or two neighboring villages that have similar living standards, in urban areas, where the distance between neighborhoods is small, enumeration areas often include communities with widely different living standards. In this study we had access to community-level data in urban areas only in the household survey.

Another reason is that in all other data sources the towns, including Ouagadougou and Bobo-Dioulasso, were considered as single points in the GIS data set. In the econometric analysis all the enumeration areas from each of the large towns have the same community characteristics and thus have to be considered as a single entity.

Of the villages in the Province of Sanguie, most of those considered poorest are located farthest from urban centers and are not connected to an all-weather road (figure 2, in the appendix). Targeting an antipoverty program to the entire population of the province is bound to include many nonpoor villages, whereas excluding this province from the program will leave out a considerable number of poor villages.

V. Simulating the Impact of a Village-Level Targeting Scheme

To evaluate the effectiveness of community targeting, we conduct a simple simulation experiment. We use the consumption data in the 201 communities for which we have complete information from the Priority Survey. The simulation design follows closely the framework of Baker and Grosh (1994). We assume that the government has a given budget for transferring income to the target population. The effect of these transfers on poverty is evaluated using the actual household consumption data. We select which villages to target, however, using the predicted levels of poverty estimated from our model. The simulations thus can evaluate how effectively these predictions identified the poor by estimating leakage and undercoverage. We make the estimates for the households included in the survey, but we use the individual sampling weights to measure the impact on the total population at the regional and national levels. The reliability of the regional and national results is therefore affected by the sampling errors that are due to the sampling frame of the Priority Survey.

We compare the outcome of this simulation with an untargeted uniform transfer scheme in which all individuals in the country receive a transfer. We also consider two other targeted programs: a village-level targeted program that uses actual poverty levels to identify the poor villages included in the program and a "perfect" targeting program that uses actual household consumption data to identify the poor villages included in the program.

The three targeted programs are designed to include 30 percent of the population.[12] To achieve this, we first set the poverty line so that 30 percent of the country's population is identified as poor. We select villages for the three targeted programs as follows. For the program that targets at the village level based on the predicted level of poverty, we rank all villages in the sample. Starting from the poorest village, we select villages for targeting until (at least) 30 percent of the population is included in the program. We follow the same procedure for the program that targets at the village level based on the actual level of poverty. For the program targeted at the household level based on actual consumption per

12. In our simulation we use a lower poverty line than in the previous section to focus on attempts to target a relatively small part of the population. Using two-thirds of mean consumption as the poverty line, poverty is estimated at 58 percent, suggesting a transfer program that attempts to include nearly two-thirds of the population. The issues of undercoverage and leakage thus become less interesting to study. The population considered in this simulation includes the households for which we have the complete set of variables in the prediction model.

adult equivalent, we again rank households, this time according to consumption level. We include households starting from the poorest until 30 percent of the population is covered.

We first consider a targeted program using actual village poverty levels.[13] The simulation results suggest that 44 percent of the poor would not be covered. With the targeted program that uses predicted village poverty levels, undercoverage would rise to 56 percent as a result of prediction errors. As an indicator of the accuracy of the predictions, this implies that about 79 percent of the poor who could be reached through targeting using actual poverty data could also be reached using predicted poverty data.

By design, untargeted transfers result in no undercoverage, but leakage is high. We define leakage as the number of nonpoor covered by a program divided by the total number of people covered. For the untargeted program leakage is (by design) 70 percent, since we consider 30 percent of the population to be poor. When using the village-level poverty estimates based on the Priority Survey data to establish criteria for targeting, leakage falls to 44 percent. The distribution of leakage across (rural) regions is similar. When using the predictions on village-level poverty to select the communities included in the scheme, leakage increases to 56 percent. Still, this amount is far less than the leakage implied by undifferentiated transfers.[14]

VI. Conclusion

Geographical targeting of antipoverty programs can be an effective way to reach the poor and contain program costs in countries where information on individual households is incomplete or unavailable, making individual or household targeting impossible. By identifying the geographic areas where the poor are concentrated, these programs can reduce leakage so that a larger share of the poor population can be reached with a given budget or a larger share of this budget can reach the poor. However, in most countries the target areas are regions, states, or entire rural areas. Although targeting at these levels can offer considerable savings compared with nontargeted programs, it invariably results in substantial leakage. Narrow targeting at the level of the community or administrative department can be an effective way to reach the poor for two main

13. Further details, including those on the interregional distribution of leakage and undercoverage, are given in Bigman and others (1999).

14. In Bigman and others (1999) we also introduce actual income transfers and evaluate the poverty impact of these transfers for a given budget. This budget is equally divided among all individuals included in each program. Using a budget that is just more than one-half the actual poverty gap, undifferentiated transfers reduce the headcount measure of poverty by a relatively high 22 percent, while village-level targeting using actual poverty levels reduces that measure of poverty by 33 percent. Using predicted poverty levels reduces the measure of poverty by about 27 percent. These are only modest gains, but giving all households living in villages included in a program the same monetary transfer is by no means optimal when minimizing poverty. Our scheme is aimed at identifying poor communities, and it attempts to minimize leakage and undercoverage.

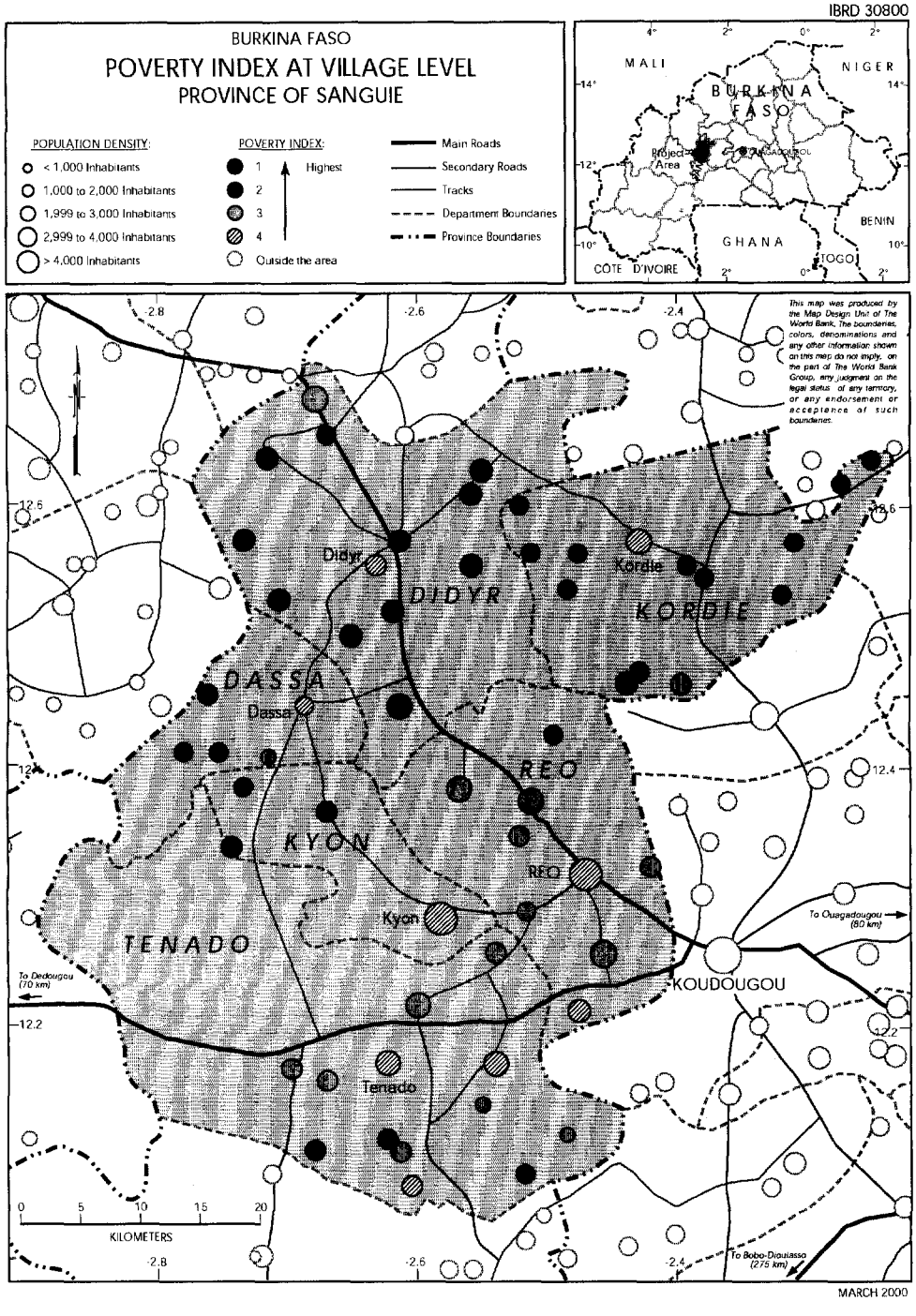
reasons. First, in most developing countries, particularly countries in Sub-Saharan Africa, poverty tends to be concentrated in villages and certain parts of towns. Second, geographical targeting entails relatively low administrative costs and, by relying primarily on local authorities, may ensure that a large portion of the benefits will reach the target population.

This article presented a methodology for using data from many different sources in order to establish criteria for targeting poverty-reduction programs at the level of the village, urban community, or local administrative department. This methodology consists of collecting data from several sources, aggregating them at the village level, and arranging them as a geographic information system. We conducted an econometric analysis using data from a household survey to identify the variables that best explain household consumption. The explanatory variables included important characteristics of the community and of households in that community. We selected the household-level explanatory variables whose mean values in each community were available for most of the communities in Burkina Faso not included in the household survey. This made it possible to use the model estimated in the regression analysis with the data of the Priority Survey to predict the incidence of poverty in all the villages outside the Priority Survey sample and thereby identify the spatial distribution of poverty at the community level.

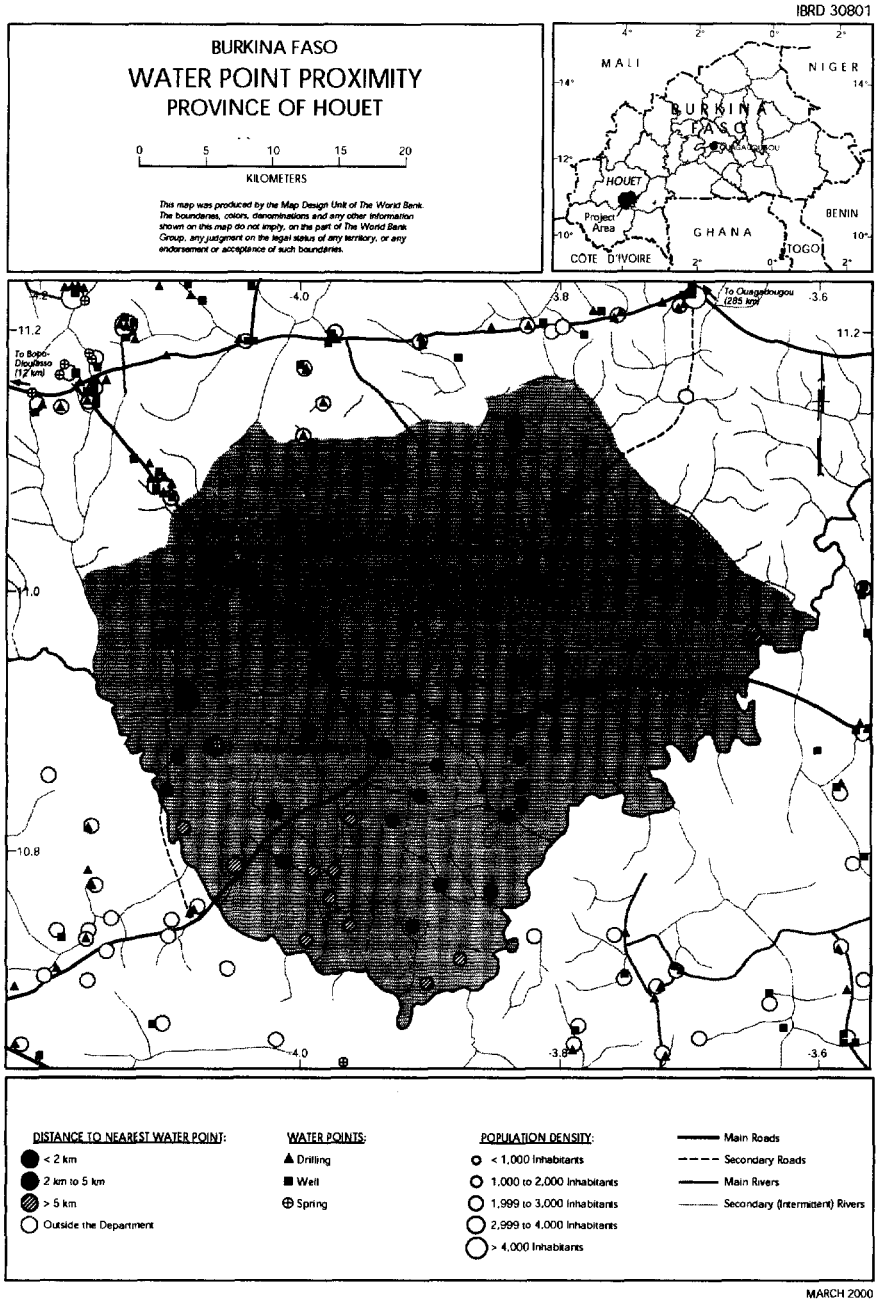
Constraints on the availability and quality of data for Burkina Faso led to considerable prediction errors and prevented us from using the complete ordering of the villages as predicted by the econometric analysis. To reduce the impact of these errors, we divided the villages into several categories and focused only on the villages categorized as poorest and least poor. Indeed, practical considerations in applying antipoverty programs and tight budget constraints are likely to reduce the need for a complete ordering. Poverty alleviation programs are more likely to focus on villages at the lower end of the distribution, and cost-recovery programs are more likely to focus on villages at the higher end. Nevertheless, the limited availability of geo-referenced data and the low quality of the data currently available reduced the predictive power of our econometric analysis. Further work is needed to augment and improve the stock of relevant data.

Targeting poverty alleviation or cost-recovery programs at the level of the village or department has other advantages as well. First, budget constraints are likely to restrict programs that are targeted to larger geographic areas, such as regions or states, and, as a result, the errors of inclusion and exclusion are likely to be high. Targeting smaller geographic areas can reach many more of the country's poor, given the same budget constraints. Second, lower-level targeting is likely to include villages and districts in all regions or states and thus be less divisive and contentious on ethnic, social, or political grounds. Third, whereas differences in the incidence of poverty among regions are primarily due to differences in agroclimatic conditions, differences in the incidence of poverty among villages within the same region often reflect past policy biases that led to differences in the quality of access roads or public services. Targeting future policies in light of these criteria can remedy past biases.

Appendix figure 1.



Appendix figure 2.



References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Baker, Judy L., and Margaret Grosh. 1994. "Measuring the Effects of Geographic Targeting on Poverty Reduction." LSMS Working Paper 99. Development Research Group, World Bank, Washington, D.C.

Besley, Tim, and Ravi Kanbur. 1991. "The Principles of Targeting." In V. N. Balasubramanyam and Sanjaya Lall, eds., *Current Issues in Development Economics*. London: Macmillan.

Bigman, David, Stefan Dercon, Dominique Guillaume, and Michel Lambotte. 1999. "Community Targeting for Poverty Reduction in Burkina Faso." CES Discussion Paper Series 99.10. Center for Economic Studies, Catholic University of Leuven. Processed.

Coulombe, Harold, and Andrew McKay. 1996. "Modeling Determinants of Poverty in Mauritania." *World Development* 24(6):1015–31.

Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Washington D.C.: World Bank and Johns Hopkins University Press.

Glesjer, Henri. 1965. "A New Test for Heteroscedasticity." *Journal of the American Statistical Association* 60(6):539–47.

Glewwe, Paul. 1991. "Investigating the Determinants of Household Welfare in Côte d'Ivoire." *Journal of Development Economics* 35(2):307–37.

Glewwe, Paul, and Oussama Kanaan. 1989. "Targeting Assistance to the Poor Using Household Survey Data." Policy, Planning, and Research Working Paper 225. Development Research Group, World Bank, Washington, D.C. Processed.

Hentschel, Jesko, Jean Olson Lanjouw, Peter Lanjouw, and Javier Poggi. 1998. "Combining Survey Data with Census Data to Construct Spatially Disaggregated Poverty Maps: A Case Study of Ecuador." Development Research Group, World Bank, Washington, D.C. Processed.

INSD (Institut National de Statistique et de Démographie). 1996. "Le profil de pauvreté au Burkina Faso." Ouagadougou. Processed.

Maddala, G. S. 1983. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge, U.K.: Cambridge University Press.

Ravallion, Martin. 1993. "Poverty Alleviation through Regional Targeting: A Case Study for Indonesia." In Karla Hoff, Avishay Braverman, and Joseph E. Stiglitz, eds., *The Economics of Rural Organization: Theory, Practice, and Policy*. New York: Oxford University Press.

———. 1996. "Issues in Measuring and Modelling Poverty." *Economic Journal* 106(September):1328–43.

Rosenzweig, Mark, and Kenneth Wolpin. 1986. "Evaluating the Effects of Optimally Distributed Public Programs." *American Economic Review* 76(3):470–82.

van de Walle, Dominique. 1995. "Public Spending and the Poor: Theory and Evidence: Introduction." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence*. Baltimore, Md.: Johns Hopkins University Press.

THE WORLD BANK ECONOMIC REVIEW, VOL. 14, NO. 1: 195-219

Combining Light Monitoring Surveys with Integrated Surveys to Improve Targeting for Poverty Reduction: The Case of Ghana

Hippolyte Fofack

Policymakers use poverty maps to design and assess poverty programs. The accuracy of these maps, which is critical for targeting, depends largely on the nature of the instrument used to construct them. Recently, in response to tight budget constraints, many countries have begun to construct poverty maps based on light monitoring surveys that rely on short questionnaires. This article shows that poverty maps constructed from such surveys are not accurate and could result in substantial leakage. Light monitoring surveys do include large samples that can help to target poverty programs at subregional levels. Combining these surveys with more detailed Integrated Surveys can help researchers reduce targeting errors significantly and build improved poverty maps with finer levels of disaggregation.

Poverty analysis and the design of targeted programs traditionally have been based on comprehensive household surveys. Such surveys are conducted infrequently, however, making it difficult to assess the effects of macroeconomic reforms on poverty and income inequality in the short and medium term. Limited budgetary resources constrain governments from conducting these surveys more often.

Policymakers fully recognize the need for reliable poverty maps, and several surveys have been designed and proposed as short-term alternatives to comprehensive surveys. These short-term instruments, which include Rapid Appraisal Methods (Narayan and Srinivasan 1994) and Priority Surveys (Marchant and Grootaert 1991), are known as light monitoring surveys. Light monitoring surveys are designed to quickly identify groups that interventions should target. In contrast, more comprehensive surveys are designed to conduct integrated poverty analysis and to draw inferences on welfare (see Ravallion 1996 and Deininger and Squire 1996). These surveys include household budget surveys, Living Standards Measurement

Hippolyte Fofack is with the Africa Region at the World Bank. His e-mail address is *hfofack@worldbank.org*. Earlier versions of this article were presented at the International Workshop on Poverty Mapping, organized by the United Nations Environment Programme and Global Resource Information Database in Norway, and at the International Conference on Geographical Targeting, organized by Johns Hopkins University in Baltimore, Md., held in October 1998. The author would like to express his warm thanks to Lionel Demery, Jack van Holst Pellekaan, and David Bigman for their useful comments and encouragement. He also is grateful to the article's anonymous referees for their valuable comments.

Study (LSMS) surveys (Grosh and Glewwe 1998), and Integrated Surveys (Delaine and others 1992). (For details on the design and implementation of these surveys, see Demery and others 1992 and Grosh and Muñoz 1996.)

Comprehensive Integrated Surveys are broad in scope, but they take a long time to administer and usually are based on relatively small samples in order to contain costs. The questionnaire collects data on most household and individual characteristics, providing extensive information on household income, credit, and savings; household enterprises; the value of durable, productive, and financial assets; agricultural livestock; food processing and consumption of own produce; food and nonfood consumption; and other expenditures. The survey is administered over the course of a year, with multiple visits to households to capture seasonality. Nonsampling errors are reduced by shortening the recall period. Using integrated surveys to target programs at the subregional level is difficult, however, because the sample sizes are too small to allow sound inferences to be drawn.

Light monitoring surveys are administered very quickly. They collect socioeconomic data on a smaller set of variables than Integrated Surveys, but their sample size is typically much larger. These surveys are based on single visits to households. They thus do not capture seasonality in consumption patterns nor provide accurate estimates of consumption and income. The short questionnaire and single visit, however, reduce the time needed for data collection and allow for a larger sample size with better representation of different geographic areas.

The limited coverage of expenditure items in light monitoring surveys has some major drawbacks for policy analysis, however. Short questionnaires focus on a few sets of goods; consumption aggregates based on these data are likely to provide estimates of total expenditures that are lower than estimates from Integrated Surveys (Deaton and Grosh forthcoming). Moreover, the underestimation of total expenditures is not likely to be uniform, shifting the Engel curve but preserving welfare rankings. Rather, the bias varies significantly across households and regions, in part because regional differences in consumption patterns and changes over time—which are determinants of the distributions of income and expenditures—are not taken into account. The distributions that emerge from the two types of surveys thus differ significantly.

Despite evidence that they might generate inaccurate estimates of aggregate expenditures, light monitoring surveys have been used extensively for policy design. In Sub-Saharan Africa, for example, World Bank poverty assessments draw major recommendations from these surveys (World Bank 1997). This article shows that welfare indicators estimated from light monitoring surveys are biased and that the bias affects not only the estimated magnitude of poverty but also its geographic distribution. The geographic distribution of poverty indicated by light monitoring surveys differs significantly from that indicated by Integrated Surveys, in part because per capita household expenditures, which are the basis for targeting, underestimate aggregate expenditures. Large differences in the level of systematic bias will yield inaccurate poverty maps; targeted income transfer schemes based on light monitoring surveys may result in leakage.

This article investigates how comprehensive Integrated Surveys can be combined with light monitoring surveys to improve geographical targeting, make transfers for poverty alleviation efficient, and sharpen the inferences on welfare measures that are drawn from light monitoring surveys. As core components of national statistical programs, light monitoring surveys and Integrated Surveys are both household-level surveys, with important similarities. Similarities in their sampling frames and sampling designs, as well as the proximity in their implementation make their combination extremely appealing for poverty and policy analysis.

I. Limitations of Light Monitoring Surveys as Instruments for Targeting

To understand the implications of using light monitoring surveys as the basis for poverty analysis, I compare welfare indicators estimated from such surveys with those derived from Integrated Surveys. To allow full comparability, I approximate aggregate expenditures as subsets of consumption items from the Integrated Survey. The study thus can be viewed as a counterfactual experiment, since estimates of total expenditures are known.

Ideally, we would like to compare an Integrated Survey and a light monitoring survey that were conducted over the same period. But the surveys were never conducted concurrently in Ghana. Moreover, in countries in which light monitoring surveys and Integrated Surveys have been carried out consecutively, variation over time and changes in the sampling frame and design make it difficult to assess the performance of these surveys from estimates of household welfare. In this study I make comparisons on households for which expenditure aggregates have been suitably constructed under the assumptions of both the full Integrated Survey and the approximated light monitoring survey. There is no time lag in data collection between the two surveys, and errors in measurement associated with variation in the sampling design are completely eliminated because the comparisons are made on the same unit of analysis.

Distribution of Expenditures from Integrated and Light Monitoring Surveys

This study is based on the third in a series of surveys of living standards in Ghana (Ghana Living Standards Survey [GLSS] 3).[1] The survey was administered to a sample of about 4,500 nationally representative households over the course of a year.[2] It collected data on all dimensions of household welfare and economic behavior, including highly disaggregated and comprehensive data on household

1. The first two surveys, conducted in 1987 and 1989, are less comprehensive than the third survey. All three surveys are based on the master sample of enumeration areas defined by the 1984 population census.

2. A multistage stratified random sampling was used in selecting the sample. Initially, 407 clusters were enumerated, and households were selected with a probability proportional to the size, with 15 households drawn in each urban cluster and 10 households in each rural cluster.

income and expenditures. Data were collected on 107 food items, and the range of nonfood items was wider than in the first two surveys (Ghana Statistical Services 1996). Estimates from the third survey are considered more accurate than estimates from the first two, in part because estimates from the third survey are based on much shorter bounded recall periods (8 recalls at two-day intervals in rural areas and 11 recalls at three-day intervals in urban areas).

Although Integrated Surveys collect extended information on both household income and expenditures, I use expenditure data as the measure of economic welfare, partly because nonsampling errors due to underreporting of income bias reported household income. There also are strong theoretical reasons for choosing expenditures over income (Deaton and Muellbauer 1980). Estimates of total household expenditures on food and nonfood constructed from the GLSS 3 data use 6 aggregates constructed from 17 subaggregates. These estimates account for all household expenses, including total household expenditures on rent (including imputed rent on owner-occupied, rent-free, or subsidized dwellings); consumption of home-produced food; and the value of wage income received by household members in the form of food. Other imputed expenditures include total wage income paid in-kind to household members, the value of produce of nonfarm enterprises consumed by households, and the use value of durable goods. The value of remittances made by households, as well as all other expenses, such as spending on education and household amenities, are also included in aggregate expenditures. Missing values and outliers are imputed on each variable based on a methodology developed by a team from the Development Economics Research Center at the University of Warwick (see Ghana Statistical Services 1996). All expenditure data are adjusted to account for inflation, which was relatively high during the survey period, using March 1992 Accra prices as the base. No adjustment is made for seasonal effects on household expenditures.

I obtain an estimate of total household expenditures by summing across all household expenditure items, subaggregates, and aggregates on food and nonfood. I first sum across items and subaggregates to obtain intermediate values for expenditure aggregates at the household level:

$$S_k^h = \sum_{j=1}^{N} \lambda \delta_j^h P_{jk}^h Q_{jk}^h \tag{1}$$

where h represents households (ranging from 1 to the total sample size), and $j = (1,2,\ldots,N)$ represents the total number of items. P_{jk}^h and Q_{jk}^h represent the price and quantity of item j, a component of aggregate k consumed by a given household h. The multiplicative factor δ_j^h is the frequency of purchase (recall period) of a given item within household h; λ is the "frequency factor" or frequency of enumerator visits to households.[3]

3. The frequency factor λ is determined from the survey design. In the GLSS 3 survey households were visited 8 times at two-day intervals in rural areas and 11 times at three-day intervals in urban areas.

A final estimate of total household expenditures for Integrated Surveys is obtained by summing across expenditure aggregates:

$$\hat{Y}_{IS}^{h} = \sum_{k=1}^{A} S_k^h = \sum_{k=1}^{A} \sum_{j=1}^{N} \lambda \delta_j^h P_{jk}^h Q_{jk}^h \tag{2}$$

where $k = (1,2,...,A)$ is the total number of subaggregates, and $\hat{Y}_{IS}^{h}$ is the total household expenditure aggregate from the Integrated Survey for household h. We obtain estimates of total household expenditures based on light monitoring surveys by similar aggregation, albeit over a much smaller number of items and subaggregates. The components of aggregate expenditures are selected following the guidelines of the standard Priority Survey, which recommends limiting the collection of expenditure data to key food and nonfood items (Demery, Grootaert, and Hill 1991). Total expenditure aggregates constructed here are based on three aggregates: two nonfood aggregates, which include expenditures on education and health, and a food aggregate, which contains 10 key food items (corn, rice, cassava, plantains, beans, groundnuts, palm oil, sugar, salt, and meat). These expenditure items, especially spending on education and food, account for a large share of total expenditures.

We can estimate total household expenditures from a Priority Survey by summing across these three aggregates, after adjusting for inflation. The adjustment for inflation is necessary because light monitoring surveys are based on a single visit to households, and inferences on welfare are drawn assuming no seasonal variation in consumption patterns—an assumption that could limit the value of these surveys for poverty analysis. An estimate of total household expenditures from the light monitoring survey is provided by:

$$\hat{Y}_{LMS}^{h} = \sum_{k=1}^{a<A} S_k^h = \sum_{k=1}^{a<A} \sum_{j=1}^{n<N} \delta_j^h P_{jk}^h Q_{jk}^h. \tag{3}$$

The frequency factor, λ, does not appear in equation 3 because the data are collected in a single visit. In the light monitoring survey both the number of aggregates and the number of items are much smaller than in the Integrated Survey ($a < A$ and $n < N$). As a result total household expenditures are generally underestimated, and $\hat{Y}_{LMS}^{h} < \hat{Y}_{IS}^{h}$. These aggregates are adjusted for household size to produce household per capita expenditures.

Summary statistics on the distributions of these adjusted variables reveal important differences (table 1). As expected, the level of welfare estimated from the light monitoring survey is much lower than that estimated from the Integrated Survey. At the national level mean per capita expenditures estimated from the light monitoring survey are about 8 percent of those estimated from the Integrated Survey. The magnitude of the difference in welfare is particularly great in rural areas. Mean per capita expenditures estimated from the light monitoring

Table 1. *Distribution of Per Capita Household Expenditures across Regions*

Agroclimatic region	Integrated Survey: Mean per capita expenditures (cedis)	Integrated Survey: Coefficient of variation	Integrated Survey: Share of national mean per capita expenditures (percent)	Light monitoring survey: Mean per capita expenditures (cedis)	Light monitoring survey: Coefficient of variation	Light monitoring survey: Share of national mean per capita expenditures (percent)	Ratio of mean per capita expenditures of light monitoring survey to Integrated Survey
Accra	260,309	155.69	1.210	38,276.16	181.30	2.165	0.1470
Other urban	225,232	155.88	1.047	33,241.86	175.10	1.880	0.1476
Rural Forest	217,835	147.40	1.012	11,628.40	231.05	0.658	0.0534
Rural Coastal	213,641	178.55	0.993	8,769.98	269.86	0.496	0.0411
Savannah	188,691	150.81	0.877	8,674.90	350.34	0.491	0.0460
All urban	233,889	157.68	1.087	34,484.26	179.13	1.951	0.1474
All rural	205,910	166.10	0.957	9,343.80	285.52	0.529	0.0454
National	215,186	163.87	1.000	17,678.70	278.50	1.000	0.0822

Source: Author's calculations based on Ghana Statistical Services (1995, 1996).

survey in urban areas are about 14 percent of those estimated from the Integrated Survey; in rural areas the light monitoring survey estimate is about 4 percent of the Integrated Survey estimate.

This bias is exacerbated by the fact that own-produced consumption, which accounts for a large share of consumption in rural areas, is not included in the light monitoring survey. The different scope of sampled items also biases the results. Increasing the number of consumption items in the light monitoring survey slightly reduces the bias between expenditure aggregates in the two surveys, but the difference remains large, and the sizable urban-rural bias persists. When the number of sampled items is increased to 20, aggregate mean per capita expenditures in the light monitoring survey rise from 8 to 10 percent of aggregate mean per capita expenditures estimated from the Integrated Survey. Increasing the number of sampled items to 30 raises the light monitoring survey estimate to 11 percent of the Integrated Survey estimate. Increasing the number of items to 20 raises the light monitoring survey estimate from 14 to 20 percent of the Integrated Survey estimate in urban areas; in rural areas the increase is only from 4 to 5 percent.[4] The persistence of a large urban-rural bias despite the increase in the number of sampled items may reflect the importance of own-produced food in rural areas.

Estimates of total household expenditures from light monitoring surveys are based on just a few items and generally are biased downward. The downward bias is likely to shift the distribution of expenditures to the left. The difference in the scope of items and subaggregates may explain the large absolute difference between the distributions revealed by the two surveys. The variation in the structure of these distributions and the large urban-rural bias may be due largely to the nature of the consumption items in the overall aggregate. Although consumption of own-production represents a large share of household consumption in rural areas, it is not accounted for in the light monitoring survey aggregate, partly because the rural economy has a low level of monetization and also because key consumption items are more tradable in urban areas, where induced transaction costs tend to increase their relative prices.

Poverty maps constructed from the light monitoring survey reveal patterns different from those constructed from the Integrated Survey. Mean per capita expenditures are much lower across all agroclimatic regions, and the size of the bias is not uniform across regions. Measured by mean per capita expenditures, Savannah is no longer the poorest region of Ghana, and the differences in mean per capita expenditures between Savannah and Rural Coastal decline. These figures represent about 50 percent of national mean per capita expenditures as estimated by the light monitoring survey. Moreover, the bias toward higher rural poverty is greater in the light monitoring survey than in the Integrated Survey. The rural mean estimate accounts for less than 30 percent of the urban mean in

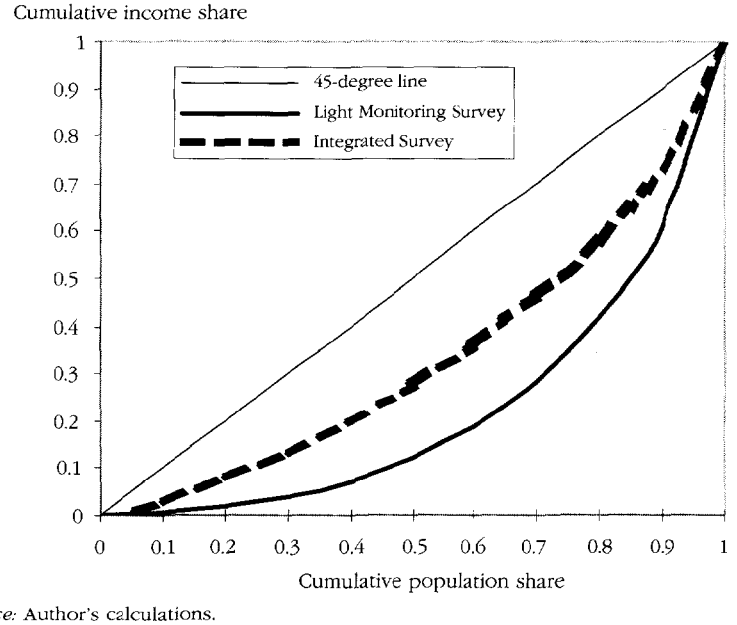
4. The national mean per capita expenditure estimates based on 20 and 30 items are 22,883 cedis and 25,803 cedis. These estimates are much higher in urban areas (44,939 cedis and 50,858 cedis) and much lower in rural areas (11,944 cedis and 13,377 cedis).

the light monitoring survey, whereas it accounts for about 90 percent of the urban mean in the Integrated Survey.

The variance in the distribution of household per capita expenditures is relatively high in the light monitoring survey estimates. Although the estimates of mean per capita expenditures are uniformly smaller in the light monitoring survey than in the Integrated Survey, the coefficient of variation of per capita expenditures is higher across all agroclimatic regions. The dispersion around the national mean is greater in the light monitoring survey, where the ratio of regional to national estimates has a larger range. Differences in mean estimates are much smaller in the Integrated Survey, where the ratio oscillates around 1, suggesting that income inequality may be much higher in the light monitoring survey. The differences in variances also are measured in terms of the Gini coefficient and the Lorenz curve of inequality (figure 1). Figure 1 supports the hypothesis that light monitoring surveys record much greater income inequality than Integrated Surveys. The Gini coefficient approximated from the light monitoring survey is substantially higher (0.56) than the corresponding coefficient for the Integrated Survey (0.34).

The relatively large variance observed in the light monitoring survey has important implications for poverty analysis. The poverty gap is proportional to the income gap, and the severity index is proportional to the squared deviation from the poverty lines when $\alpha \geq 1$. And, unexpectedly, large variances are likely to increase the estimate of the poverty gap (and therefore the cost of poverty reduction, which is proportional to the income gap).

Figure 1. *Lorenz Curve of Income Inequality*



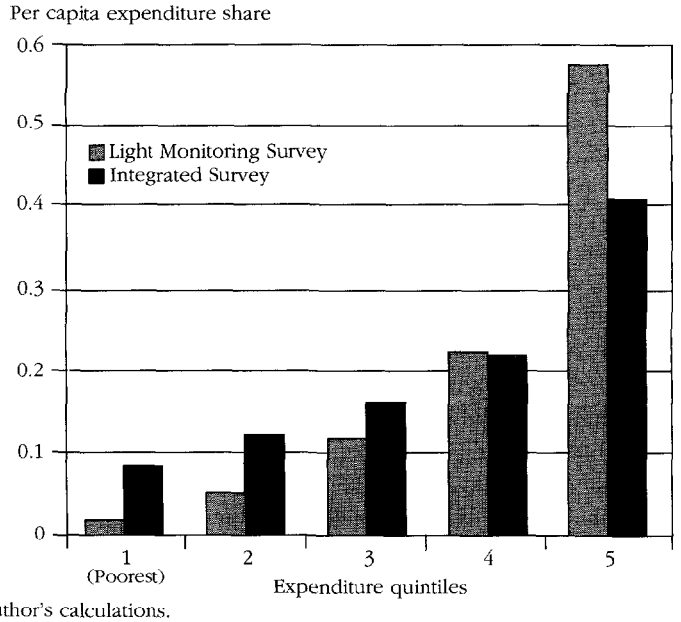
Source: Author's calculations.

These large variations are also illustrated by the distribution of per capita household expenditures across expenditure quintiles, which differs significantly between the two surveys.[5] The bottom 40 percent of the population accounts for less than 8 percent of total expenditures in the light monitoring survey and 20 percent of all expenditures in the Integrated Survey. In the upper quintiles, expenditure shares are overestimated in the light monitoring survey (figure 2). These differences in the distribution of expenditures across quintiles suggest that poverty maps constructed from a light monitoring survey may not be accurate because the instrument tends to overestimate the welfare of the nonpoor and underestimate the welfare of the poor.

Implications for Policy Analysis and Poverty Mapping

The latest Ghana poverty profile uses two measures of poverty (Boateng and others 1990). The broader measure includes all people whose per capita expenditures are two-thirds or less of the national mean (equal to 132,230 cedis per person per year).[6] The narrower measure of poverty includes all people whose per capita expenditures amount to no more than half of the national mean (equal

Figure 2. *Distribution of Per Capita Household Expenditures*



Source: Author's calculations.

5. Expenditure quintiles for the Integrated Survey are derived from the full consumption aggregate adjusted for household size; expenditure quintiles for the light monitoring survey are constructed from consumption aggregates adjusted for household size, but with a limited number of items.

6. The same definitions were used in the 1987–88 poverty profile. The 1988 base values were adjusted for inflation and expressed in 1992 constant prices for the 1995 profile. The lower poverty line is equal to the same share of national mean per capita expenditures estimated from the full GLSS 3.

to 107,188 cedis per person per year). I use the lower poverty line (z_{IS} = 107,188) as the cutoff point in the Integrated Survey, because maximum targeting is easily achieved at the lower and upper ends of the distribution, where the within-group variance and the probability of household misclassification are lower than in the middle of the distribution. Similarly, I define the relative poverty line to be half of mean per capita expenditures as approximated by the light monitoring survey (z_{LMS} = 8,839). I use that measure for cross-sectional analysis, looking at variations in poverty rates across regions in the two surveys during the same period. I do so because total household expenditures aggregated from the light monitoring survey show a large urban-rural bias—the rural expenditure aggregate is substantially lower—and because using an upper poverty line would exacerbate the scope of rural poverty.

In order to assess the performance of the light monitoring survey as a targeting instrument, I estimate the headcount, poverty gap, and severity indexes from distributions in both the light monitoring survey and the Integrated Survey (table 2).[7] Performance is assessed by the probability of type I and type II errors, as well as by the rate of mistargeting.

The probability of a type I error is defined formally as $\varepsilon_I = [P(y_j \in P \mid y_j \in \bar{P})]$, where P represents the set of poor households or individuals (y_j) and $\bar{P}$ represents the set of nonpoor households or individuals. A type I error can be referred to as an error of inclusion because it indicates the probability of classifying nonpoor households or individuals as poor.

The probability of a type II error is defined as $\varepsilon_{II} = [P(y_j \in \bar{P} \mid y_j \in P)]$. A type II error can be referred to as an error of exclusion because it gives the probability of classifying poor households or individuals as nonpoor. The rate of mistargeting depends on the size of these two errors. Perfect targeting is achieved when the rate of mistargeting is equal to 1.0, implying that both surveys classify the same number of individuals as poor. Perfect targeting occurs when the errors of inclusion and exclusion are both close to 0. Let $\zeta_{RM}(\varepsilon_I, \varepsilon_{II})$ be the rate of mistargeting expressed as a function of the error of inclusion (ε_I) and the error of exclusion (ε_{II}). This rate is a number between 0 and n, where $n < \infty$, that is, $0 < \zeta_{RM}(\varepsilon_I, \varepsilon_{II}) < n$. When mistargeting results largely from the error of inclusion, $\zeta_{RM}(\varepsilon_I, \varepsilon_{II}) > 1$. When the error of exclusion is much larger than the error of inclusion, the rate of mistargeting is confined between 0 and 1, that is, $0 < \zeta_{RM}(\varepsilon_I, \varepsilon_{II}) < 1$.

When estimates of aggregate per capita household expenditures from the light monitoring survey are used as the basis for constructing poverty maps, Rural Coastal and Savannah remain the poorest regions in Ghana. The magnitudes of the differences across regions indicated by the two surveys vary significantly, however. In GLSS 3 urban expenditures exceed rural expenditures by just 4 per-

7. The welfare indexes are selected from the P_α class of poverty indexes (Foster, Greer, and Thorbecke 1984), which measures different dimensions of poverty depending on the value of α. When $\alpha = 0$, the P_α represents the headcount index; when $\alpha = 1$, it measures the poverty gap index. The indexes provide estimates of the severity of poverty when $\alpha > 1$.

Table 2. *Indexes of Extreme Poverty and Rate of Mistargeting across Regions*

	Light monitoring survey			*Integrated Survey*			*Light monitoring survey*		
Agroclimatic region	*Headcount index*	*Poverty gap index*	*Severity index*	*Headcount index*	*Poverty gap index*	*Severity index*	*Type I error probability*	*Type II error probability*	*Rate of mistargeting*
Accra	0.0678	0.0274	0.0158	0.1429	0.0259	0.0072	0.0369	0.1121	0.47
Other urban	0.0865	0.0325	0.0180	0.1657	0.0350	0.0118	0.0423	0.1214	0.52
Rural Forest	0.5543	0.3029	0.2060	0.1673	0.0324	0.0097	0.3978	0.0108	3.31
Rural Coastal	0.6731	0.3722	0.2493	0.1887	0.0409	0.0135	0.4889	0.0045	3.57
Savannah	0.6810	0.4236	0.3082	0.2471	0.0546	0.0173	0.4752	0.0414	2.76
All urban	0.0819	0.0313	0.0175	0.1600	0.0327	0.0106	0.0410	0.1191	0.51
All rural	0.6506	0.3753	0.2604	0.2043	0.0438	0.0141	0.4648	0.0186	3.18
National	0.4621	0.2612	0.1799	0.1897	0.0401	0.0129	0.3243	0.0519	2.44

Source: Author's calculations based on Ghana Statistical Services (1995, 1996).

centage points. In contrast, the light monitoring survey indicates a difference of 55 percent. Similarly, the light monitoring survey classifies more than 68 percent of the population of the Savannah region as extremely poor (falling below the lower poverty line), while the Integrated Survey classifies just 25 percent of the population as poor. The same variations are observed in other regions, with poverty overestimated in rural areas and underestimated in urban areas.

A poverty map constructed from the GLSS 3 data has a higher urban headcount than a poverty map constructed from light monitoring survey data, partly because the GLSS 3 accounts for consumption of own production in rural areas, which reduces the urban-rural bias and increases national mean per capita expenditures (and therefore the extreme poverty line). Thus the light monitoring survey underestimates the scope of urban poverty. However, in the comprehensive Integrated Survey the headcount index is much higher in urban areas, and targeting may be justified as well.

Poverty in Sub-Saharan Africa is generally much higher in rural areas, where opportunities for income generation are much more limited than in urban areas. Extreme differences such as those revealed by the light monitoring survey—which classifies more than 65 percent of the rural population and less than 8 percent of the urban population as living in extreme poverty—are nevertheless unexpected. These large urban-rural differences reflect differences in the error of inclusion, which is much higher in rural areas than in urban areas (0.47 compared with 0.04). The size of the error is directly proportional to the rate of misclassification; the larger is the error of inclusion, the higher is the rate of misclassification.

In urban areas the error of inclusion is low, and the error of exclusion is much higher (0.11). Mistargeting results from large errors of exclusion, which occur because poor households are underrepresented in the sample of intended beneficiaries. Large errors of inclusion in rural areas result from underestimating total household consumption and overrepresenting the poor population in the approximated light monitoring surveys, which causes nonpoor households to be surveyed as intended beneficiaries for targeted interventions. Mistargeting is relatively high under the light monitoring survey design, in which the population identified for the targeted intervention is nearly two and a half times larger than the true population estimate (table 2). This high rate of overall mistargeting is inflated by the rural rate of misclassification, which is more than three times the true number of intended beneficiaries. The variations in the rates of mistargeting across other rural regions (Rural Forest, Rural Coastal, and Savannah) are not significant.

The amount of leakage is directly proportional to the rate of misclassification. It will be lower in urban areas, where differential rates are smaller. The number of extremely poor people mistargeted by the light monitoring survey is about half the true targeted population in urban areas; in rural areas the number of misclassified people is three times the number of primary beneficiaries. The dollar amount of leakage that occurs as a result of poor targeting by the light monitoring survey is more than three times the amount required to alleviate extreme poverty in rural Ghana (table 2).

The estimated cost of eradicating extreme poverty ($nz\hat{p}_1$) is proportional to the poverty gap, where $\hat{p}_1$ is the estimate of the poverty gap. The poverty gap estimated from the light monitoring survey is about six times higher than the GLSS estimate at the national level. As a result, more than six times as much money would be needed to eradicate extreme poverty if the light monitoring survey were used as the basis for poverty analysis (27,998 cedis per capita per year compared with 4,298.5 cedis).[8] The potential costs to the central government and local authorities are considerable, because poor targeting and improper identification of intended beneficiaries increase the amount of leakage and the amount of resources allocated for poverty alleviation.

II. Predicting Consumption Expenditures to Improve Targeting

The accuracy of light monitoring survey data can be improved by using poverty predictors, correlates of expenditures that are used to impute household consumption. Fofack (1997) proposes a methodology for deriving national poverty predictors that could be used to improve targeting. Poverty predictors and their corresponding weights are estimated from Integrated Surveys and used to predict total household expenditures. These predicted values then serve as the basis for conducting poverty analysis and for constructing poverty maps from light monitoring survey data. The prediction error is low, and high rates of successful classification are achieved. Moreover, a test of stability shows that the poverty predictors and their corresponding weights are stable over time.[9]

Recently, attempts have been made to exploit the wide coverage of population censuses by combining them with household surveys (Hentschel and others 1998). Although combining these tools is appealing—especially given the scope for geographical targeting—the data requirements for capturing the large proportional variance in welfare could be enormous. The method proposed by Hentschel and others uses a large number of regressors from the census to predict welfare. Here, I take a different approach, based on data reduction. The poverty correlates used to predict welfare are reduced to a set of minimum core variables that can be easily collected with minimal measurement error.

To model household consumption for poverty analysis, the best correlates of welfare are derived from the GLSS 3 survey using correlation analysis and regression models. The model assumes that the conditional expectation of the response, given the covariates, $E(y|x_1,...,x_k)$, is related to the linear predictors by the re-

8. This is a hypothetical scenario included for illustrative purposes only. Direct income transfers are not cornerstones of policy in Sub-Saharan Africa. The cost of such measures would be enormous and would worsen fiscal deficits.

9. The stability of the poverty predictors and the corresponding weights estimated from GLSS 3 (1992) are assessed by applying the regressors from GLSS 3 to GLSS 1 (1987) and GLSS 2 (1989) to predict the level of welfare in those years. Success rates of ranking households in the same expenditure quintiles are as high as 95 percent, despite the time lag between these surveys. For details on the implementation of the test of stability of these regressors, see Fofack (1997).

sponse link function $h(x,\theta)$. Because the variance of total household expenditures across and within regions is large, a logarithmic transformation is applied to the response to make the relationship between y and the x's linear. This transformation stabilizes the error variance, reduces asymmetry in the distribution of error terms, and improves the prediction. The structural form of the correcting model is specified by equation 4:

$$Y = \mathbf{X}'\boldsymbol{\beta} + \varepsilon \tag{4}$$

where Y is total household expenditures transformed to the log scale, $\mathbf{X}$ is the model matrix containing the vector of regressors, $\boldsymbol{\beta}$ is the vector of estimated parameters relative to continuous and discrete-level variables, and ε are the error terms, which are distributed as $N(0,\sigma^2)$. The poverty predictors are predominantly discrete-level variables. Most continuous variables with strong predictive capabilities are dichotomized to discriminate between poor and nonpoor households. These dummy regressors are constructed and included in the model to capture the effects of qualitative independent variables.

In order to account for selection bias in choosing predictor variables, I use a conditional maximum likelihood estimation method to select predictors. Unlike other selection criteria, the conditional maximum likelihood method is based on the expected overall discrepancy. Because the omission bias in the fixed model becomes an additional residual variation, the method produces an unbiased estimator of the discrepancy.[10] The best poverty predictors are those that significantly increase the explanatory power of the model. That is, if $(x_1,x_2,\ldots,x_j)$ is the initial set of poverty correlates and x_{j+1}, x_{j+k} for $k \neq 1$ are potential poverty predictors, the variable x_{j+1} will be selected over x_{j+k} if

$$\Sigma[y_j - E(\hat{y} \mid x_1,x_2,\ldots,x_j,x_{j+1})]^2 < \Sigma[y_j - E(\hat{y} \mid x_1,x_2,\ldots,x_j,x_{j+k})]^2 \tag{5}$$

Initially, I assume that all predictor variables are available for inclusion in the model. I then proceed by elimination, using the stepwise selection method with a minimum level of significance. I remove a given independent variable from the model only when a marginal increase in the percentage variance of the response explained by the model as a result of that variable's inclusion is smaller than the marginal increase attributed to the inclusion of any other independent variable:

$$\Delta\left[R^2\left(y \,\middle|\, \sum_{j=1}^{k} \lambda_j x_j\right) - R^2\left(y \,\middle|\, \sum_{j=1}^{k-1} \lambda_j x_j\right)\right] < \varepsilon. \tag{6}$$

Applying this selection procedure to the model iteratively produces an optimal model with 10 core poverty predictors. The optimal model has very few continu-

10. Other criteria used to select the subset of predictors include the S_p and the Mallows C_p criteria. The S_p is a model selection criterion that assumes that the response variable and the predictors are jointly normally distributed. The C_p method, due to Mallows (1973), assumes that the predictor variables are fixed and not random. For more details on model selection see Linhart and Zucchini (1986).

ous variables; those that remain are either dichotomized or discrete-level variables. This characteristic is likely to reduce errors that arise because of long recall periods. Moreover, the accuracy of targeting increases because the poverty predictors and the weighted coefficients are estimated from Integrated Survey data and are imputed using information collected at the household level during the administration of the light monitoring survey. The poverty indexes thus are no longer just a function of aggregate household expenditures, but also depend on the estimated regression coefficients:

$$P_k = f(\hat{y}, z), \text{ for } k = 0, 1, 2, \tag{7}$$

where $\hat{y} = \Phi^{-1}\left(\sum_{j=1}^{n} \beta_j x_j\right)$, and z is the poverty line.

The poverty predictors are derived at the national and regional levels and are used in conjunction with the corresponding weights to predict total expenditures. Predicted expenditures, expressed as the weighted sum of the poverty predictors, are then used as the basis for constructing poverty maps, classifying regions for poverty analysis, and targeting.

The poverty predictors are able to explain more than 65 percent of the proportional variance observed in the welfare measure reported. The proportional variance explained by the model is high at the national level and at the regional level when the model is calibrated to derive poverty predictors for each agroclimatic region. (The appendix lists the derived poverty predictors at the national and regional levels. The use of different poverty predictors is intended to reflect differences in consumption patterns.)

III. Empirical Results

To assess the accuracy of poverty maps constructed from the improved light monitoring survey—that is, the light monitoring survey in which total expenditures have been modeled using poverty predictors—I estimate the incidence of poverty and poverty-related indicators for different agroclimatic regions using predicted expenditures constructed from the model. I compare these estimates with the poverty indicators derived from GLSS 3 using the same poverty line for the predicted and measured consumption aggregates (since in the absence of prediction errors, the means of these two distributions are the same).

The differences in the poverty estimates decrease substantially in both urban and rural areas when the poverty predictors are used to model household expenditures (table 3). The error of inclusion, for example, which was as high as 0.48 in the approximated light monitoring survey, falls to just 0.13 or less when the poverty predictors are the basis for poverty analysis. Moreover, the significant decrease in this error is not accompanied by an increase in the error of exclusion, which remains low across all agroclimatic regions.

Table 3. *Indexes of Extreme Poverty and Rate of Mistargeting by Region with the Improved Light Monitoring Survey*

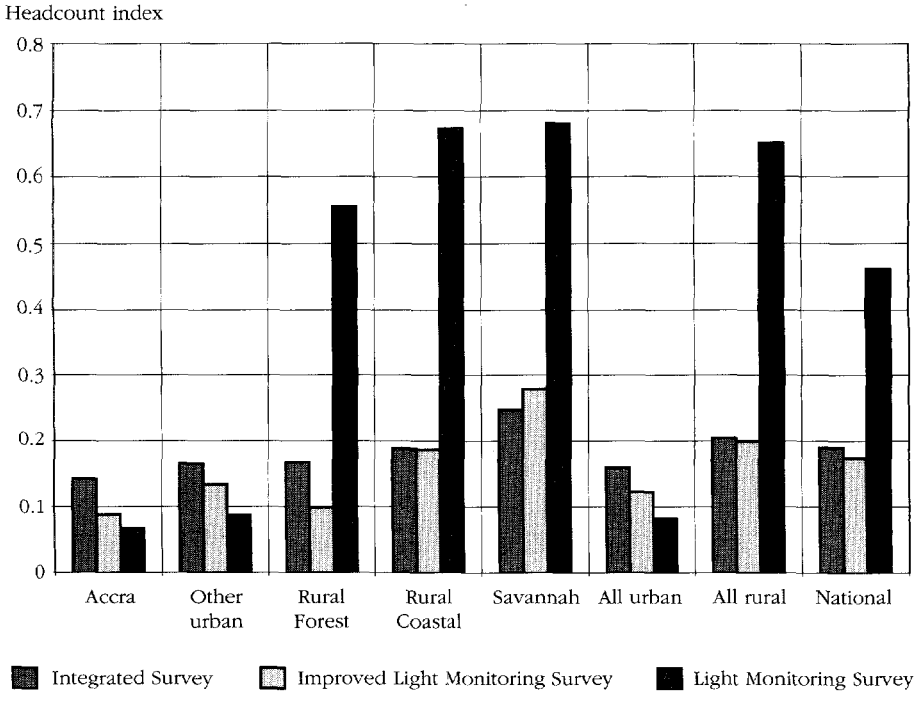
	Improved light monitoring survey			*Integrated Survey*			*Improved light monitoring survey*		
Agroclimatic region	*Headcount index*	*Poverty gap index*	*Severity index*	*Headcount index*	*Poverty gap index*	*Severity index*	*Type I error probability*	*Type II error probability*	*Rate of mistargeting*
Accra	0.0878	0.0163	0.0040	0.1429	0.0259	0.0072	0.0212	0.0763	0.615
Other urban	0.1344	0.0274	0.0076	0.1657	0.0350	0.0118	0.0623	0.0936	0.811
Rural Forest	0.0988	0.0159	0.0039	0.1673	0.0324	0.0097	0.0537	0.1220	0.591
Rural Coastal	0.1868	0.0357	0.0103	0.1887	0.0409	0.0135	0.1036	0.1056	0.990
Savannah	0.2787	0.0461	0.0116	0.2471	0.0546	0.0173	0.1261	0.0946	1.128
All urban	0.1229	0.0247	0.0067	0.1600	0.0327	0.0106	0.0522	0.0894	0.768
All rural	0.1999	0.0351	0.0094	0.2043	0.0438	0.0141	0.1008	0.1053	0.978
National	0.1743	0.0317	0.0085	0.1897	0.0401	0.0129	0.0847	0.1001	0.919

Source: Author's calculations based on Ghana Statistical Services (1995, 1996).

The poverty indicators estimated on the basis of the GLSS 3 data and the improved light monitoring survey data are similar (figure 3). At the national level the absolute relative error is less than 0.081, probably because the difference in the headcount indexes estimated from the Integrated Survey and the predicted welfare function is small. This relatively small difference is due largely to measurement errors in reported household expenditures and sample size effects. The sample size at the subregional level is the smallest in Accra and the Rural Forest region, where the magnitude of the difference is largest.

The prediction error is inversely proportional to the sample size, however, suggesting that the error should decrease substantially in the actual light monitoring survey, where the large sample size and exhaustive coverage allow much greater representation at the regional and subregional levels. Part of the prediction error also could be attributed to measurement errors in reported household expenditures. This part of the prediction error is associated with the survey design and implementation and could be large depending on sample size and nonsampling errors. In contrast, measurement errors are washed away by the instrumented variable; targeting might even become more accurate when based on predicted welfare. A potential increase in accuracy is another reason for using instrumented consumption in poverty analysis.

Figure 3. *Headcount Index*



Source: Author's calculations.

The differences between headcount indexes remain when predicted expenditures are used for poverty analysis. The bias toward higher rural poverty is preserved in the regional ranking, and the ranking of the five agroclimatic regions is consistent across the two ranking criteria (Integrated Survey data and improved light monitoring survey data with predicted expenditures). The magnitude of the differences across agroclimatic regions is also preserved in the improved light monitoring survey.

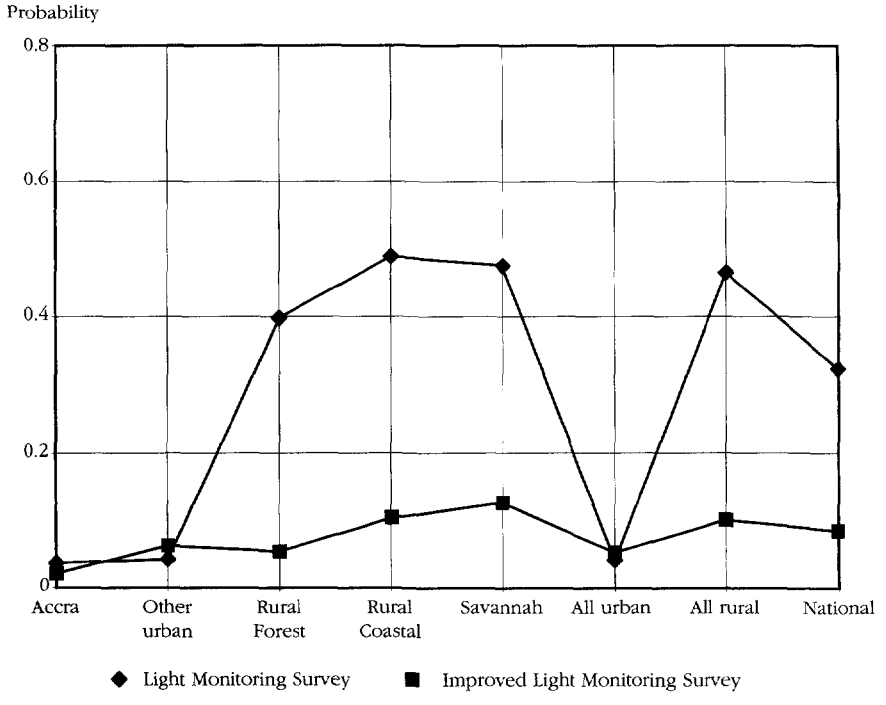
The incidence of poverty is highest in the Savannah and Rural Coastal regions, where 28 and 19 percent of the population, respectively, are extremely poor. The largest poverty gap is in Savannah, suggesting that if the population were distributed uniformly across regions, the volume of transfers needed to eradicate poverty would be largest in this region. The Savannah and Rural Coastal regions also have the lowest literacy and enrollment rates, and both depend on food crops as their main source of income (Ghana Statistical Services 1995). In contrast, in the Rural Forest region, where income sources are more diversified, the incidence of poverty is much lower. The poorest of the poor are thus uneducated small farmers residing mainly in the Savannah and Rural Coastal regions, where employment opportunities are limited. To the extent that education variables (public school enrollment, proportion of school-age children enrolled in school) appear to be good proxies for income in these regions, the placement of public infrastructure may be a good means of making transfers to the poor.

I also assess the performance of the proposed method by the size of errors of inclusion and exclusion, as well as the rate of mistargeting. While the error of inclusion across agroclimatic regions is generally much higher in the approximated light monitoring survey than in the improved light monitoring survey (figure 4), the error of exclusion is generally much lower (figure 5). The magnitude of the difference between the errors of exclusion in the two surveys is small, however, in part because mistargeting is due largely to high errors of inclusion.

The rate of mistargeting falls substantially when the improved light monitoring survey is used (table 3). In fact, except in the Rural Forest region, targeting is almost perfect. Perfect targeting is achieved in all Rural and Rural Coastal areas when predicted expenditures are used as the basis for constructing poverty maps. Compared with the light monitoring survey, the gains in accuracy are significant. Even in the Rural Forest region, where the error of inclusion is relatively high (0.12) and the rate of misclassification is slightly different from unity, the gap is less than 40 percent. Moreover, the rate of mistargeting is less than 1, implying that poor targeting is due largely to a high error of exclusion.

The rates of mistargeting $\zeta_{RM}(\varepsilon_I, \varepsilon_{II})$ are particularly high in rural areas when the approximated light monitoring survey is used to construct the poverty map. These high rates are attributable mostly to large errors of inclusion. Significant improvement is achieved when predicted expenditures are used as the basis for targeting (in the improved light monitoring survey).

Figure 4. *Probability of Error of Inclusion*



Source: Author's calculations

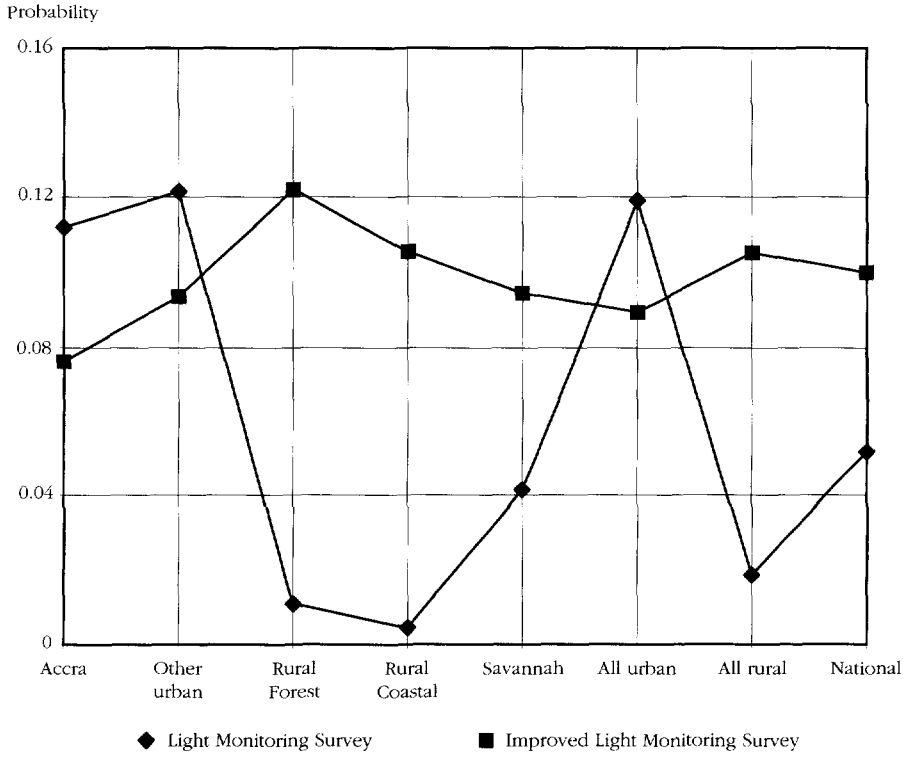
IV. Implications for Geographical Targeting

The method described sharpens the accuracy of poverty maps and allows policymakers to target beneficiaries at subregional levels. The approximated light monitoring survey is constructed from the more comprehensive GLSS 3 survey. The sample size corresponding to the approximated light monitoring survey is dictated by the GLSS 3 design, just as the level of disaggregation is determined by the actual GLSS 3 sample size.

The geographic profile of poverty provides living standards indicators for the various agroclimatic regions. Efficiency could be improved and leakage reduced significantly if smaller geographical units could be targeted (Baker and Grosh 1994). The Priority Survey design recommends a large sample size for targeting smaller administrative units (Grootaert and Marchant 1992).[11] Improving the accuracy of the welfare function by predicting household expenditures should enable researchers to exploit the large sample size to achieve geographical target-

11. The Kenya Welfare Monitoring Survey (1994) was based on a sample of 12,000 households. The Ghana Core Welfare Indicators Survey (1997) was based on a sample of 15,000 households.

Figure 5. *Probability of Error of Exclusion*



Source: Author's calculations.

ing with minimum leakage at a level of disaggregation well below the agroclimatic region.

The causes and determinants of poverty, as well as the sources of large disparities across agroclimatic regions, are variable. At the aggregate level differences in the potential for income-generating activities and wage inequality may constitute important factors; at the regional and district levels human capital, access indicators, and location of infrastructure may be more critical. Light monitoring surveys collect good data on access indicators (location of schools, health centers, and water supply). The relatively large sample size of such surveys may provide opportunities for geo-referencing information at subregional levels, thereby improving the potential for analysis beyond fixed geographical boundaries. Further, overlaying improved poverty maps atop maps of local infrastructure (schools, health clinics, hospitals, water supply facilities, and roads) may improve the understanding of poverty dynamics, shed more light on the possible constraints to growth and poverty reduction, and improve priority setting, impact assessment, and policymaking.

Poverty predictors, which include food and nonfood consumption variables, are strong correlates of welfare. They can serve as the basis for targeting by commodity and by welfare indicator, especially if indirect transfer schemes are used to reduce poverty.[12] Targeting by indicator is based on the ability to easily identify a few key variables that are highly correlated with household income and expenditures. The use of regional poverty predictors, which are correlates of welfare, may be particularly appealing in Sub-Saharan Africa, where there are important differences in the determinants of welfare across agroclimatic regions. Taking these differences into account in selecting correlates might improve the efficiency of transfers and the allocation of public expenditures for poverty reduction.[13]

The poverty predictors also can serve as a basis for targeting by commodity because they include food items. This targeted scheme draws on the differences observed in the consumption baskets of the poor and nonpoor. Its objective is to reduce the cost of commodities that are consumed heavily by the poor through targeted subsidies. Although poverty predictors are derived according to agroclimatic region rather than poverty, the methodology is flexible and could be used in multiple steps—that is, poverty predictors could first be used to predict household expenditures, and expenditures in turn could be used to differentiate between the poor and the nonpoor. A cross-sectional analysis that focuses on the variation in the consumption of the poor by agroclimatic region could be a starting point for investigating the causal link between variation in the depth of poverty and the nature of poverty correlates. Future research will have to explore the links between these correlates and poverty at the regional and district levels, and determine how a better understanding of those associations could be used to channel scarce resources to the most needy.

V. Concluding Remarks

Many developing countries are confronted with widespread poverty and have limited resources for poverty alleviation. To minimize leakage, policymakers must have accurate and detailed poverty maps that allow identification of the poor at finer levels of disaggregation than the agroclimatic region.

Some developing countries have used light monitoring surveys, which have large sample sizes and are less expensive to administer than other types of surveys, as the basis for constructing disaggregated poverty maps. This study shows that the cost of mistargeting associated with the use of such surveys is significant and can outweigh the savings generated from their lower administrative costs. Light monitoring surveys underestimate aggregate expenditures, the basis for dif-

12. Other methods of targeting include targeting by income and self-targeting. For a survey of targeting methods with applications to developing countries, see Glewwe (1992); Kanbur, Keen, and Tuomala (1994); Bigman and Fofack (forthcoming).

13. Glewwe (1992) uses housing indicators as the basis for targeting in Côte d'Ivoire. Ravallion (1989) uses landownership as the basis for designing land-contingent transfers for poverty alleviation in Bangladesh.

ferentiating between the poor and the nonpoor. Moreover, underestimation is not uniform across regions. As a result, welfare indicators and poverty maps derived from light monitoring surveys may not always be consistent with the actual distribution of poverty.

This article shows that combining more detailed surveys, which have comprehensive income and expenditure data, with light monitoring surveys yields improved poverty maps that are disaggregated at a level below the agroclimatic region. The rate of mistargeting is reduced substantially when poverty predictors derived from more comprehensive surveys are used to model total expenditures based on data from light monitoring surveys. These poverty predictors are household-level variables that are available in both Integrated Surveys and light monitoring surveys.

Over the past few years demand for poverty maps that are disaggregated at a level as low as the district has been growing in developing countries, particularly in Sub-Saharan Africa. This demand has been prompted both by the need for a more accurate picture of the geographic distribution of poverty and by the move toward decentralization, which increasingly channels resources to communities. As the demand for more disaggregated information continues to grow and budgetary and resource constraints tighten, methods that optimize the use of light monitoring surveys while improving the accuracy of targeting will be in increasing demand. The method proposed here recommends that light monitoring surveys and more detailed comprehensive surveys be combined to improve poverty mapping and geographical targeting.

The accuracy of household welfare predicted from the model depends on the base point of the prediction, the stability of the poverty predictors, and their corresponding weights. Although modeling consumption significantly reduces errors of inclusion and exclusion, the level of targeting attained in the various agroclimatic regions is imperfect because of unavoidable prediction errors. These errors are attributable largely to measurement errors in reported expenditures and might be washed away when the instrumented variable is used.

The stability of the predictors over time is another important question. Fofack (1997) assesses stability using surveys conducted under the same sampling frame. It would be worth investigating how this stability is affected by variation in the sampling frame. The possible implications for poverty mapping are also worth examining. Finally, efficiency in the transfer and allocation of resources could be improved by combining geographical targeting with another form of targeting, such as targeting by commodity or by indicator. Poverty predictors are correlates of expenditures and could serve as a vector of transfers if the dynamic between these correlates and poverty were better understood.

Appendix. *National and Regional Poverty Predictors*

National level	*Urban areas*	*Rural areas*
Expenditures on soap	Expenditures on soap	Expenditures on soap
Number of spouses	Number of spouses	Number of spouses
Asset score	Asset score	Asset score
Percentage of school-age children enrolled in school	Percentage of school-age children enrolled in school	Percentage of school-age children enrolled in school
Expenditures on meat	Expenditures on meat	Expenditures on meat
Ownership of land	Ownership of land	Ownership of land
Consumption of bread	Percentage of household members employed	Ownership of goats and sheep
Ownership of poultry	Use of toothpaste	Number of household members per room
Export crops	Percentage of children enrolled in public school	Ownership of farm
Number of household members per room	Percentage of literate household members	Ownership of cattle
Accra region	*Other urban*	*Rural Forest*
Asset score	Expenditures on soap	Expenditures on soap
Expenditures on meat	Number of household members employed	Consumption of bread
Percentage of household members employed	Expenditures on meat	Asset score
Expenditures on rice	Asset score	Use of toothpaste
Number of household members who completed secondary school	Percentage of school-age children	Number of spouses
Expenditures on soap	Expenditures on bread	Expenditures on meat
Percentage of school-age children	Percentage of household members who completed secondary school	Percentage of household members who completed secondary school
Number of children under five	Use of toothpaste	Expenditures on rice
Use of toilet paper	Ownership of land	Number of household members per room
Percentage of children enrolled in public school	Use of toilet paper	Use of toilet paper
Rural Coastal	*Savannah*	
Expenditures on soap	Expenditures on soap	
Asset score	Number of spouses	
Percentage of school-age children	Consumption of bread	
Expenditures on meat	Percentage of school-age children	
Number of spouses	Ownership of sheep and goats	
Percentage of children enrolled in public school	Use of toothpaste	
Consumption of bread	Expenditures on meat	
Use of toilet paper	Asset score	
Ownership of poultry	Use of toilet paper	
Number of children under five	Gender of head	

REFERENCES

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Baker, Judy L., and Margaret Grosh. 1994. "Measuring the Effects of Geographical Targeting on Poverty Reduction." Living Standards Measurement Study Working Paper 99, World Bank, Washington, D.C.

Bigman, David, and Hippolyte Fofack. Forthcoming. "An Overview of Targeted Schemes for Poverty Reduction in Developing Countries." In David Bigman and Hippolyte Fofack, eds., *Geographical Targeting for Poverty Alleviation: Methodology and Applications*. Washington, D.C.: World Bank.

Boateng, E. Oti, Kodwo Ewusi, Ravi Kanbur, and Andrew D. McKay. 1990. *A Poverty Profile for Ghana, 1987–1988*. Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper 5. Washington, D.C.: World Bank.

Deaton, Angus, and Margaret Grosh. Forthcoming. "Consumption." In Margaret Grosh and Paul Glewwe, eds., *Designing Household Survey Questionnaires: Lessons from Ten Years of LSMS Experience for Developing Countries*. Washington, D.C.: World Bank.

Deaton, Angus, and John Muellbauer. 1980. *Economics of Consumer Behavior*. Cambridge, U.K.: Cambridge University Press.

Deininger, Klaus, and Lyn Squire. 1996. "A New Data Set Measuring Income and Inequality." *The World Bank Economic Review* 10(September):565–91.

Delaine, Ghislaine, Lionel Demery, Jean-Luc Dubois, Branko Grdjic, Christiaan Grootaert, Christopher Hill, Timothy Marchant, Andrew McKay, Jeffrey Round, and Christopher Scott, eds. 1992. *The Social Dimensions of Adjustment Integrated Survey: A Survey to Measure Poverty and Understand the Effects of Policy Change on Households*. Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper 14. Washington, D.C.: World Bank.

Demery, Lionel, Christiaan Grootaert, and Christopher Hill. 1991. "Annotated Questionnaire and Listing Forms." In Timothy J. Marchant and Christiaan Grootaert, eds., *The Social Dimensions of Adjustment Priority Survey: An Instrument for the Rapid Identification and Monitoring of Policy Target Groups*. Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper 12. Washington, D.C.: World Bank.

Demery, Lionel, Jean-Luc Dubois, Christiaan Grootaert, and Tim Marchant. 1992. "Annotated Integrated Survey Questionnaire." In Ghislaine Delaine, Lionel Demery, Jean-Luc Dubois, Branko Grdjic, Christiaan Grootaen, Christopher Hill, Timothy Marchant, Andrew McKay, Jeffrey Round, and Christopher Scott, eds., *The Social Dimensions of Adjustment Integrated Survey: A Survey to Measure Poverty and Understand the Effects of Policy Change on Households*. Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper 14. Washington, D.C.: World Bank.

Fofack, Hippolyte. 1997. "Using Poverty Predictors as Expenditure Proxies for Ranking Households for Poverty Analysis." *Proceedings of the International Statistical Institute 1997*. Ankara: State Institute of Statistics.

Foster, James E., Joel Greer, and Erik Thorbecke. 1984. "A Class of Decomposable Poverty Measures." *Econometrica* 52(3):173–77.

Ghana Statistical Services. 1995. "Ghana Living Standards Survey Report on the Third Round (GLSS 3)." Accra.

———. 1996. "Measuring Household Income and Expenditure in the Third Round of the Ghana Living Standards Survey (1991/92): A Methodological Guide." Accra.

Glewwe, Paul. 1992. "Targeting Assistance to the Poor: Efficient Allocation of Transfers When Household Income Is Not Observed." *Journal of Development Economics* 38(2):297–321.

Grootaert, Christiaan, and Timothy J. Marchant. 1992. "SDA Socioeconomic Information System." In Ghislaine Delaine, Lionel Demery, Jean-Luc Dubois, Branko Grdjic, Christiaan Grootaen, Christopher Hill, Timothy Marchant, Andrew McKay, Jeffrey Round, and Christopher Scott, eds., *The Social Dimensions of Adjustment Integrated Surveys: A Survey to Measure Poverty and Understand the Effects of Policy Change on Households.* Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper 14. Washington, D.C.: World Bank.

Grosh, Margaret, and Paul Glewwe. 1998. "Data Watch: The World Bank's Living Standards Measurement Household Surveys." *Journal of Economic Perspectives* 12(1):187–96.

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

Hentschel, Jesko, Jean Olson Lanjouw, Peter Lanjouw, and Javier Poggi. 1998. "Combining Census and Survey Data to Study Spatial Dimensions of Poverty." Policy Research Working Paper 1928. Development Research Group and the Poverty Reduction and Economic Management Network, World Bank, Washington, D.C. Processed.

Kanbur, Ravi, Michael Keen, and Matti Tuomala. 1994. "Labor Supply and Targeting in Poverty Alleviation Programs." *The World Bank Economic Review* 8(2):191–211.

Linhart, Heinz, and Walter Zucchini. 1986. *Model Selection.* Wiley Series in Probability and Mathematical Statistics. New York: John Wiley.

Mallows, C. L. 1973. "Some Comments on C_p." *Technometrics* 15(4):661–75.

Marchant, Timothy, and Christiaan Grootaert, eds. 1991. *The Social Dimensions of Adjustment Priority Survey: An Instrument for the Rapid Identification and Monitoring of Policy Target Groups.* Social Dimensions of Adjustment in Sub-Saharan Africa Working Paper Working Paper 12. Washington, D.C.: World Bank.

Narayan, Deepa, and Lyra Srinivasan. 1994. *Participatory Development Tool Kit: Materials to Facilitate Community Empowerment.* Washington, D.C.: World Bank.

Ravallion, Martin. 1989. "Land-Contingent Poverty Alleviation Scheme." *World Development* 17(8):1223–33.

———. 1996. "How Well Can Method Substitute for Data? Five Experiments in Poverty Analysis." *The World Bank Research Observer* 11(2):199–221.

World Bank. 1997. "Status Report on Poverty in Sub-Saharan Africa 1997: Tracking the Incidence and Characteristics of Poverty." Institutional and Social Policy Department, Africa Region, World Bank, Washington, D.C. Processed.

Coming in the next issue of

THE WORLD BANK ECONOMIC REVIEW

May 2000
Volume 14, Number 2

The World Bank

1818 H Street, N.W.
Washington, D.C. 20433, U.S.A.
Telephone: (202) 477-1234
Facsimile: (202) 477-6391
Telex: MCI 64145 WORLDBANK or MCI 248423 WORLDBANK
World Wide Web: http://www.worldbank.org/
E-mail: books@worldbank.org

14607
9 780821 346075

ISBN 0-8213-4607-5