



THE WORLD BANK

# Research Observer

---

EDITOR

Shantayanan Devarajan, World Bank

CO-EDITOR

Gershon Feder, World Bank

EDITORIAL BOARD

Susan Collins, Georgetown University

Angus Deaton, Princeton University

Barry Eichengreen, University of California–Berkeley

Emmanuel Jimenez, World Bank

Benno Ndulu, World Bank

Howard Pack, University of Pennsylvania

Luis Servén, World Bank

Sudhir Shetty, World Bank

Michael Walton, World Bank

*The World Bank Research Observer* is intended for anyone who has a professional interest in development. *Observer* articles are written to be accessible to nonspecialist readers; contributors examine key issues in development economics, survey the literature and the latest World Bank research, and debate issues of development policy. Articles are reviewed by an editorial board drawn from across the Bank and the international community of economists. Inconsistency with Bank policy is not grounds for rejection.

The journal welcomes editorial comments and responses, which will be considered for publication to the extent that space permits. On occasion the *Observer* considers unsolicited contributions. Any reader interested in preparing such an article is invited to submit a proposal of not more than two pages to the Editor. Please direct all editorial correspondence to the Editor, *The World Bank Research Observer*, 1818 H Street, NW, Washington, DC 20433. USA.

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 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.

---

For more information, please visit the Web sites of the **Research Observer** at  
[www.wbro.oxfordjournals.org](http://www.wbro.oxfordjournals.org), the World Bank at [www.worldbank.org](http://www.worldbank.org),  
and Oxford University Press at [www.oxfordjournals.org](http://www.oxfordjournals.org).

# Research Observer

---

Volume 21 • Number 2 • Fall 2006

Macroeconomic Stability in Developing Countries: How Much Is Enough? Peter Montiel and Luis Servén	151
Banking and Regulation in Emerging Markets: The Role of External Discipline Xavier Vives	179
Access to Financial Services: A Review of the Issues and Public Policy Objectives Stijn Claessens	207
What Have We Learned from a Decade of Manufacturing Enterprise Surveys in Africa? Arne Bigsten and Måns Soderbom	241
Is There a Case for Industrial Policy? A Critical Survey Howard Pack and Kamal Saggi	267

## Subscriptions

A subscription to *The World Bank Research Observer* (ISSN 0257-3032) comprises 2 issues. Prices include postage: for subscribers outside the Americas, issues are sent air freight.

### Annual Subscription Rate (Volume 21, 2 issues, 2006)

Academic libraries

Print edition and site-wide online access: US\$115/£73/€110

Print edition only: US\$109/£69/€104

Site-wide online access only: US\$104/£66/€99

Corporate

Print edition and site-wide online access: US\$172/£109/€164

Print edition only: US\$164/£104/€156

Site-wide online access only: US\$155/£98/€147

Personal

Print edition and individual online access: US\$41/£29/€44

Please note: £ Sterling rates apply in the UK, € in Europe, US\$ elsewhere.

Readers with mailing addresses in non-OECD countries and in socialist economies in transition are eligible to receive complimentary subscriptions on request by writing to the *UK address* below.

There may be other subscription rates available; for a complete listing, please visit [www.wbro.oxfordjournals.org/subscriptions](http://www.wbro.oxfordjournals.org/subscriptions).

Full prepayment in the correct currency is required for all orders. Orders are regarded as firm, and payments are not refundable. Subscriptions are accepted and entered on a complete volume basis. **Claims** cannot be considered more than four months after publication or date of order, whichever is later. All subscriptions in Canada are subject to GST. Subscriptions in the EU may be subject to European VAT. If registered, please supply details to avoid unnecessary charges. For subscriptions that include online versions, a proportion of the subscription price may be subject to UK VAT. Personal rates are applicable only when a subscription is for individual use and are not available if delivery is made to a corporate address.

**Back issues:** The current year and two previous years' issues are available from Oxford University Press. Previous volumes can be obtained from the Periodicals Service Company, 11 Main Street, Germantown, NY 12526, USA. E-mail: [psc@periodicals.com](mailto:psc@periodicals.com). Tel: (518) 537-4700. Fax: (518) 537-5899.

**Contact information:** Journals Customer Service Department, Oxford University Press, Great Clarendon Street, Oxford OX2 6DP, UK. E-mail: [jnls.cust.serv@oxfordjournals.org](mailto:jnls.cust.serv@oxfordjournals.org). Tel: +44 (0)1865 353907. Fax: +44 (0)1865 353485.

**In the Americas, please contact:** Journals Customer Service Department, Oxford University Press, 2001 Evans Road, Cary, NC 27513, USA. E-mail: [jnlororders@oxfordjournals.org](mailto:jnlororders@oxfordjournals.org). Tel: (800) 852-7323 (toll-free in USA/Canada) or (919) 677-0977. Fax: (919) 677-1714. **In Japan, please contact:** Journals Customer Service Department, Oxford University Press, 1-1-17-5F, Mukogoaoka, Bunkyo-ku, Tokyo, 113-0023, Japan. E-mail: [okudaoup@po.ijnet.or.jp](mailto:okudaoup@po.ijnet.or.jp). Tel: (03) 3813 1461. Fax: (03) 3818 1522.

**Postal information:** *The World Bank Research Observer* (ISSN 0257-3032) is published by Oxford University Press for the International Bank for Reconstruction and Development/*THE WORLD BANK*. Send address changes to *The World Bank Research Observer*, Journals Customer Service Department, Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009. Communications regarding original articles and editorial management should be addressed to The Editor, *The World Bank Research Observer*, *The World Bank*, 1818 H Street, NW, Washington, DC 20433, USA.

**Permissions:** For information on how to request permissions to reproduce articles or information from this journal, please visit [www.oxfordjournals.org/jnls/permissions](http://www.oxfordjournals.org/jnls/permissions).

**Advertising:** Inquiries about advertising should be sent to Helen Pearson, Oxford Journals Advertising, PO Box 347, Abingdon OX14 1GJ, UK. E-mail: [helen@oxfordads.com](mailto:helen@oxfordads.com). Tel: +44 (0)1235 201904. Fax: +44 (0)8704 296864.

**Disclaimer:** Statements of fact and opinion in the articles in *The World Bank Research Observer* are those of the respective authors and contributors and not of the International Bank for Reconstruction and Development/*THE WORLD BANK* or Oxford University Press. Neither Oxford University Press nor the International Bank for Reconstruction and Development/*THE WORLD BANK* make any representation, express or implied, in respect of the accuracy of the material in this journal and cannot accept any legal responsibility or liability for any errors or omissions that may be made. The reader should make her or his own evaluation as to the appropriateness or otherwise of any experimental technique described.

**Paper used:** *The World Bank Research Observer* is printed on acid-free paper that meets the minimum requirements of ANSI Standard Z39.48-1984 (Permanence of Paper).

**Indexing and abstracting:** *The World Bank Research Observer* is indexed and/or abstracted by *ABI/INFORM*, CAB Abstracts, Current Contents/Social and Behavioral Sciences, Journal of Economic Literature/*EconLit*, *PAIS International*, *RePEc* (Research in Economic Papers), Social Services Citation Index, and Wilson Business Abstracts.

**Copyright** © The International Bank for Reconstruction and Development/*THE WORLD BANK*

All rights reserved; no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise without prior written permission of the publisher or a license permitting restricted copying issued in the UK by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1P 9HE, or in the USA by the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

---

# Macroeconomic Stability in Developing Countries: How Much Is Enough?

---

Peter Montiel • Luis Servén

---

*Over the 1990s macroeconomic policies improved in most developing countries, but the growth dividend from this improvement fell short of expectations, and a policy agenda focused on stability turned out to be associated with a multiplicity of financial crises. This article examines the contents and implementation of the macroeconomic reform agenda of the 1990s. It reviews the progress achieved through fiscal, monetary, and exchange rate policies across the developing world and the effectiveness of the changing policy framework in promoting stability and growth. The main lesson is that more often than not slow growth and frequent crises resulted from shortcomings in the reform agenda of the 1990s. These concern limitations in the depth and scope of the reform agenda, its lack of attention to macroeconomic vulnerabilities, and its inadequate attention to complementary reforms outside the macroeconomic sphere.*

For developing countries the 1990s were characterized by two major macroeconomic developments: improvements in macroeconomic policies and a proliferation of financial crises. Although macroeconomic policies as traditionally measured improved in most countries, the growth benefits expected from these better policies failed to materialize—at least to the extent anticipated by many observers—and a series of financial crises had adverse effects on economic growth and poverty in the countries involved. This article examines the relationship between these two developments. It argues that slow growth and multiple crises were symptoms of deficiencies in the design and execution of the pro-growth reform strategy adopted in the 1990s, of which macroeconomic stability was viewed as the centerpiece.<sup>1</sup>

A useful way to characterize the interpretation of recent growth experience proposed here is from the perspective of Rodrik (2004) "growth strategies." He argued that well-established property rights, market-oriented incentives, fiscal solvency, and price stability are first-order economic principles that are necessary conditions for rapid economic growth. He stressed that these conditions can be implemented

through a variety of institutional arrangements and identified two views on whether establishing these conditions is also sufficient to accelerate economic growth. One view suggests that rapid growth is simply waiting to happen once the conditions are met; the other contends that more proactive government policies are also required. This evaluation of macroeconomic reform's role in growth takes the second perspective. It has four parts.

First, improvements in fiscal solvency and price stability did not occur in all developing countries. As a consequence, macroeconomic instability continued to impede growth in some countries, and in several cases traditional macroeconomic imbalances resulted in crises during the 1990s that resembled those of the 1980s.

Second, what matters for growth is the private sector's perception that fiscal solvency and price stability will be sustainable, which requires underlying fiscal and monetary rules and institutions to be reformed. Improved macroeconomic policy realizations were much more widespread than reforms in the rules and institutions governing macroeconomic policy formation. The limited progress in reforming macroeconomic institutions likely undermined the contribution of macroeconomic policy improvements—even when they could have been sustained *ex post*—to the stimulation of economic growth.

Third, the macroeconomic stability sought through fiscal solvency and price stability was undermined by fragility caused by misguided reform policies in the domestic financial system and the capital account of the balance of payments. These policies left many stabilizing economies highly vulnerable to adverse shocks and proved to be the Achilles' heel of macroeconomic stability in some of the most important crises of the 1990s.

Fourth, and perhaps most important, the growth payoff of macroeconomic stability *per se* may have been oversold. Fiscal solvency and price stability are conducive for growth because macroeconomic instability hampers investors' ability and willingness to undertake investment opportunities—understood in the broadest sense of the term. But for macroeconomic stability to deliver growth, those opportunities must exist in the first place. In other words, macroeconomic stability may not be the binding constraint that prevents accelerated economic growth. Not only are well-defined property rights and market-oriented incentives also important for satisfactory growth, but if the proactive government view mentioned above is correct, a variety of other measures may also be required. In short, macroeconomic stability can help, but by itself it cannot deliver growth. Unfortunately, gains in macroeconomic stability were often not complemented by the necessary growth-enhancing reforms in other parts of the economy.

The rest of this article develops these arguments by examining the macroeconomic reform agenda of the 1990s. It first reviews progress in implementing the reform agenda during the past decade. It then evaluates the effectiveness of the reforms from an economic growth perspective and discusses how a policy agenda

focused on macroeconomic stability was associated with a multiplicity of financial crises. It concludes by summarizing the lessons from the experience of the 1990s.

## The Facts of the 1990s

Macroeconomic instability refers to phenomena that make the domestic macroeconomic environment less predictable. Unpredictability hampers resource allocation decisions, investment, and growth.<sup>2</sup> It can take the form of volatile key macroeconomic variables or perceived unsustainability in their behavior. This section evaluates developing countries' gains in macroeconomic stability during the 1990s, looking separately at the behavior of macroeconomic outcome variables, policy variables, and exogenous shocks.

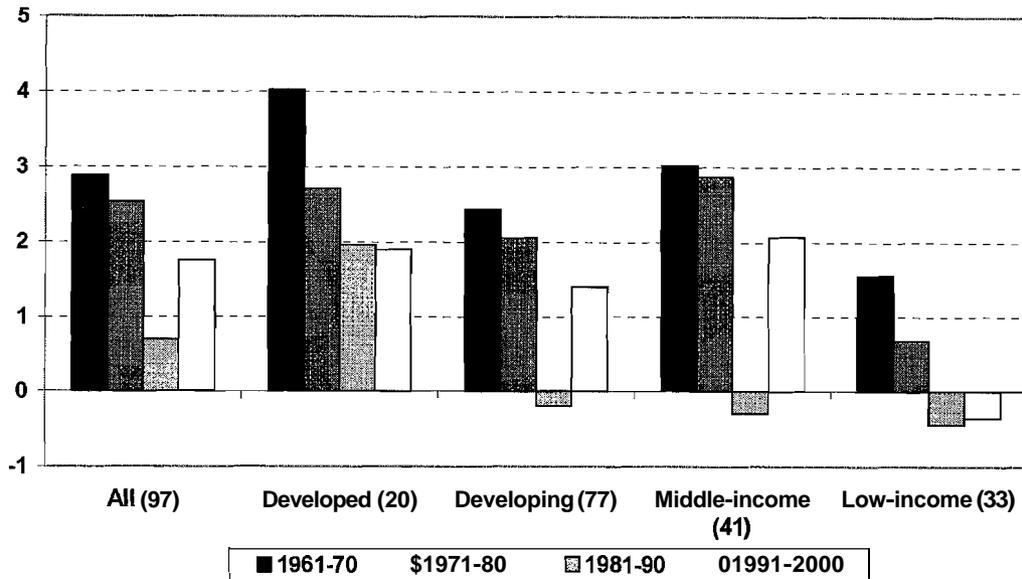
### *Macroeconomic Outcomes*

For developing countries, growth rebounded in the 1990s from the depths of the 1980s, but it still fell far short of the levels achieved in the late 1960s and 1970s and lagged behind the growth of industrial countries (figure 1). Low-income countries did much worse than middle-income countries in the 1990s, showing little improvement relative to the 1980s, whereas middle-income countries' growth rates in the 1990s were roughly on par with those of developed economies: much higher than in the 1980s but well below those in the 1960s and 1970s.

What about the stability of growth outcomes? Developing countries have traditionally been characterized by more macroeconomic instability than developed economies, and there is a widespread perception that globalization has made the situation worse (IADB 1995; De Ferranti and others 2000; Easterly, Islam, and Stiglitz 2001; Rodrik 2001). However, the volatility of key macroeconomic aggregates actually declined in the 1990s across the developing world. The standard deviation of per capita GDP growth fell from 4 percent in the 1970s and 1980s to about 3 percent in the 1990s, although it still remained significantly above the 1.5 percent seen in developed economies (figure 2).<sup>3</sup> Although the reduction in volatility of GDP growth was widespread, it was far from universal; of the 77 developing countries for which complete data are available over 1960–2000, more than a third (27 countries) experienced an increase in growth volatility in the 1990s relative to the 1980s.<sup>4</sup>

Moreover, the reduction in aggregate output volatility concealed increasing extreme instability. Large growth disturbances accounted for a larger share of overall instability in the 1990s than in previous decades because of the increased contribution of large negative shocks (crises), which accounted for close to 25 percent of total growth volatility compared with 14 percent in the 1960s and 1970s and 18 percent

**Figure 1. Median Real GDP Growth, by Decade and Country Income Group (Percent)**



Note: The sample comprises 97 countries with a population above 500,000 that have complete data on real GDP growth over 1960–2000. The population minimum is set to exclude highly volatile island economies. Of the 77 developing economies, three (Hong Kong, China, Israel, and Singapore) are high-income, non-Organisation for Economic Co-operation and Development economies.

Source: World Bank (various years); Hnatkovska and Loayza (2004).

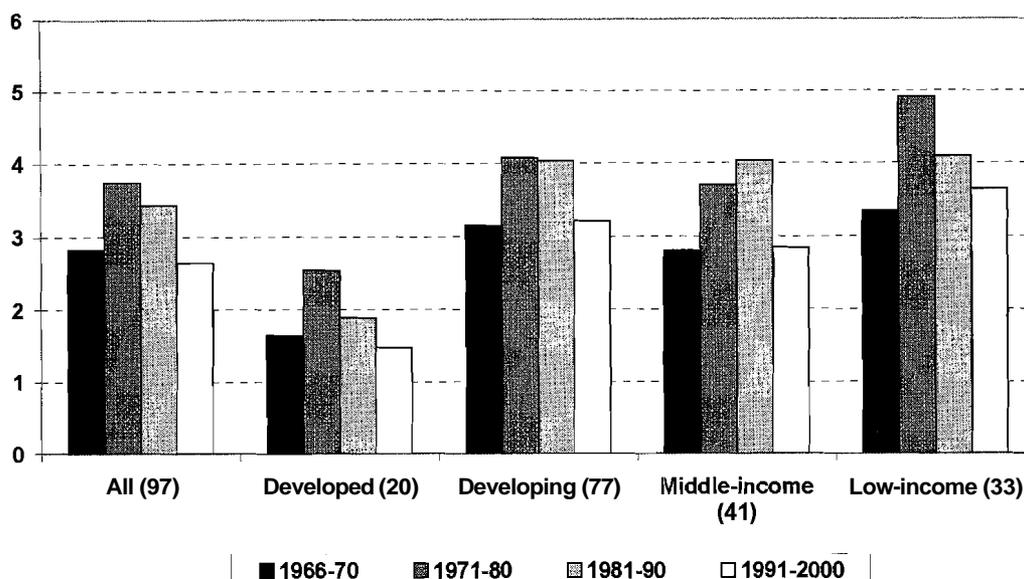
in the 1980s (figure 3). Negative extreme shocks also accounted for a larger share of the total volatility of gross national income and consumption in the 1990s than in previous decades.

Other key outcome variables commonly used as indicators of macroeconomic stability improved in the 1990s. For example, the median inflation rate across middle-income countries declined from a peak of 16 percent in 1990 to 6 percent in 2000. In low-income countries inflation peaked in 1994–95, after the devaluation of the CFA franc, and then declined (figure 4). Yet over most of the 1990s the gap between developed and developing country median inflation rates was substantial by the standards of the 1960s and 1970s.

Likewise, the prevalence of high inflation in developing countries peaked in 1991 and then declined sharply. However, the decline took hold only in the mid-1990s, and thus the share of developing countries (among those with complete data) experiencing average inflation above 50 percent over the decade as a whole was unchanged between the 1980s and the early 1990s.

Finally, current account deficits followed disparate trends in low- and middle-income countries. In middle-income countries the median current account deficit to GDP

**Figure 2.** Standard Deviation of per Capita GDP Growth, by Decade and Country Income Group (Percent)



*Note:* The sample comprises 97 countries with a population above 500,000 that have complete data on real GDP growth over 1960–2000. The population minimum is set to exclude highly volatile island economies. Of the 77 developing economies, three (Hong Kong, China, Israel, and Singapore) are high-income, non-Organisation for Economic Co-operation and Development economies.

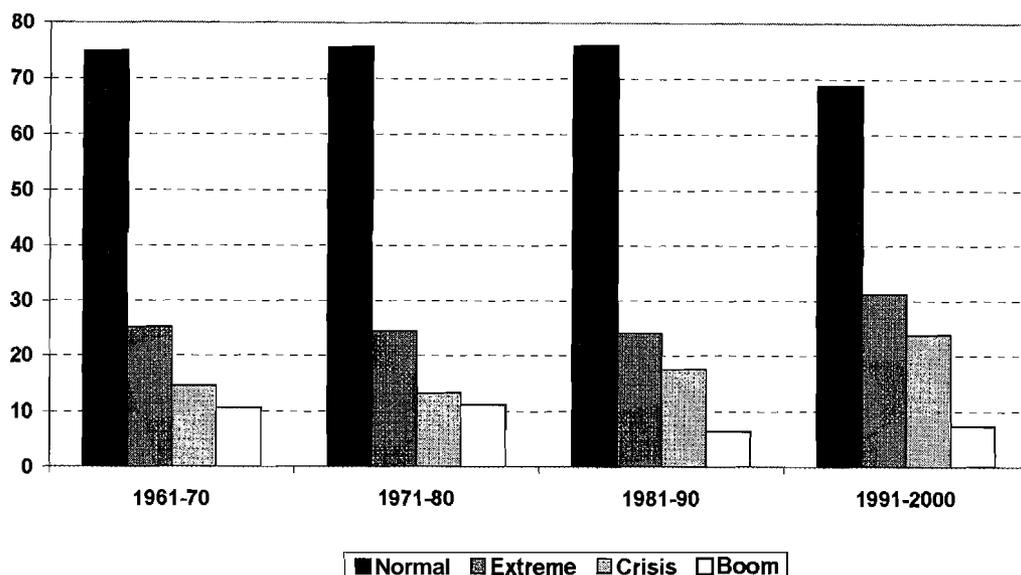
*Source:* World Bank (various years); Hnatkovska and Loayza (2004).

ratio fell by about 1 percentage point from the 1970s and 1980s. In part, however, this apparent improvement reflects the sudden stop of capital inflows to crisis-afflicted emerging market economies. In low-income countries, the deficit rose by about half a point to almost 5 percent in the 1990s (figure 5).

### *Policy Stability*

Conventional indicators of policy stability also show a broad improvement over the 1990s. Most notably, the overall fiscal deficit fell across the developing world from a median value of 6–7 percent of GDP in the early 1980s to 2 percent in the early 1990s before rebounding to about 3 percent by the end of the decade. The fiscal correction was particularly pronounced among middle-income countries. However, the overall fiscal balance is affected by the impact of interest rate changes on public debt, which is beyond the direct control of the authorities. Thus the primary balance likely offers a more accurate measure of fiscal stance. Over the 1990s it shows a clear trend of increasing surpluses, particularly after 1995 (figure 6). By the end of the

**Figure 3.** Decomposition of Growth Volatility in Developing Countries, by Decade (Mean Percentage of Total Volatility)



*Note:* Total volatility = normal + extreme; extreme = crisis + boom. Extreme shocks are defined as those exceeding two standard deviations of output growth over the respective decade.

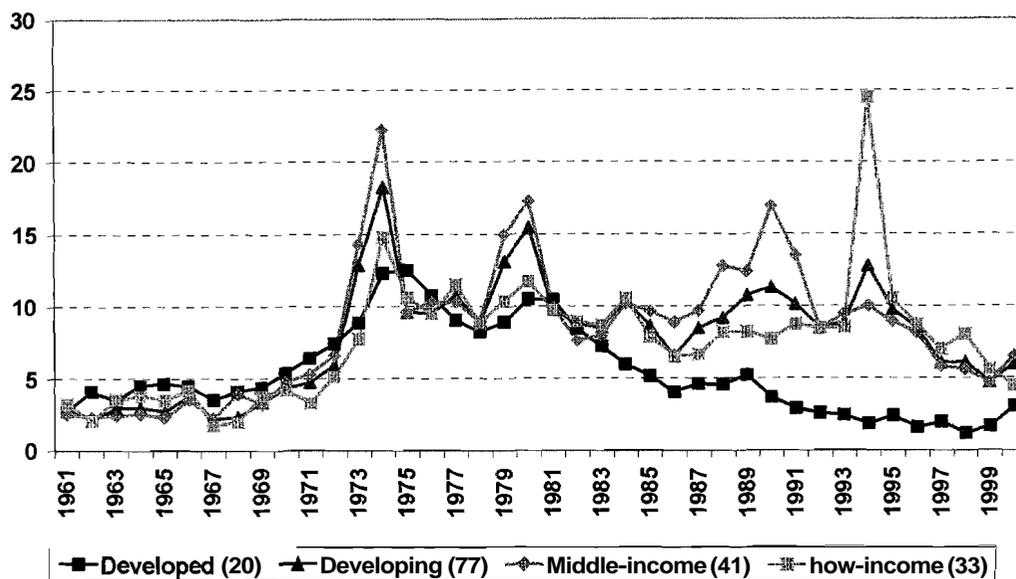
*Source:* Authors' calculations based on data from Hnatkovska and Loayza (2004).

decade the median developing country showed a primary surplus—although a much more modest one than that of industrial countries.

Because of the diversity of monetary arrangements across developing countries and over time, it is more difficult to gauge monetary stability. One rough measure is the resort to seigniorage—that is, using money to finance deficits. Measured by the change in the monetary base relative to GDP, seigniorage collection rose during the late 1980s and early 1990s and then declined in both middle-income and (more modestly) low-income economies (figure 7), a pattern roughly similar to that of the inflation rate.

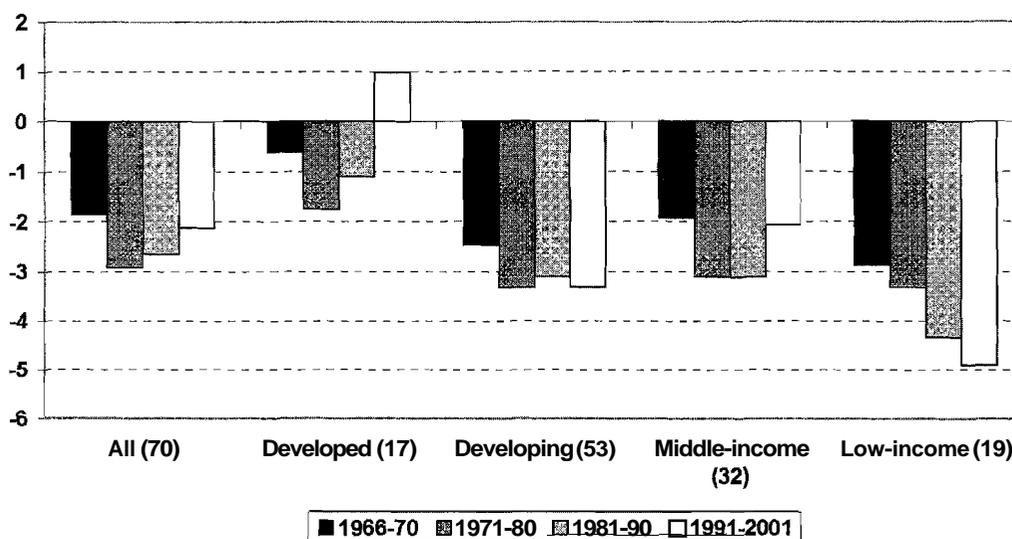
The diversity of exchange rate arrangements across countries also makes it hard to gauge exchange rate policy. One indirect approach looks at the trends in real exchange rates, which are of course endogenous and subject to the influence of a variety of factors, including the nominal exchange rate. Shvets (2004) showed that real exchange rates depreciated over the 1990s in most developing economies. At the same time real exchange rate volatility (as measured by the standard deviation of its rate of change) showed a decline from the record high levels of the 1980s. But the decline was limited to middle-income countries, and over the 1990s developing countries as a group exhibited much higher real exchange rate volatility than

**Figure 4.** Median Inflation Rate, by Country Income Group, 1961–2000 (GDP Deflator)



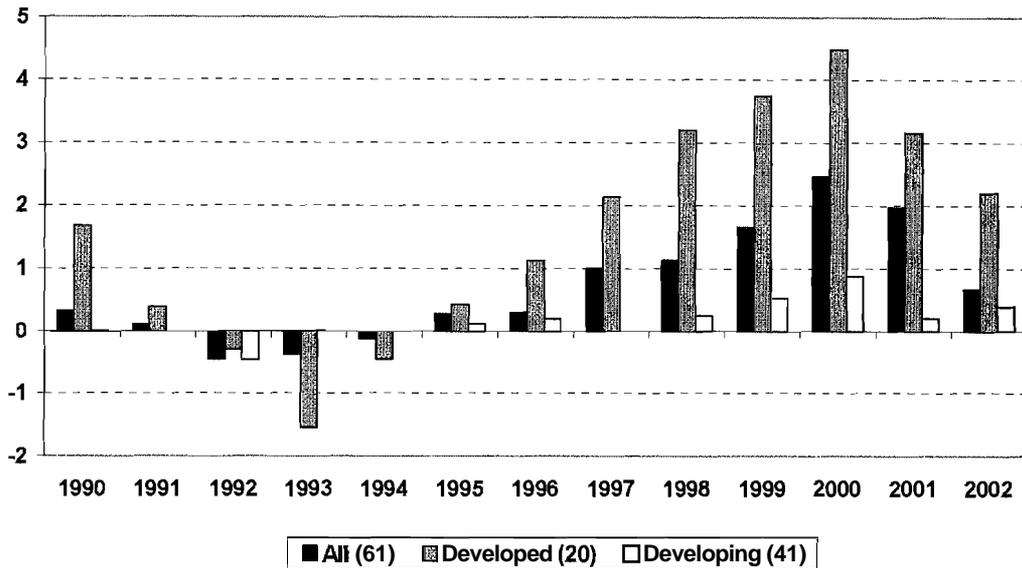
Source: World Bank (various years).

**Figure 5.** Median Current Account Balance, by Decade and Country Income Group (Percent of GDP)



Source: World Bank (various years).

**Figure 6.** Median Primary Fiscal Balance, by Country Income Group, 1990–2002 (Percent of GDP)



*Note:* Data differ in source and coverage from those in figure 7; the figures are not strictly comparable.  
*Source:* Fitch Ratings (various years).

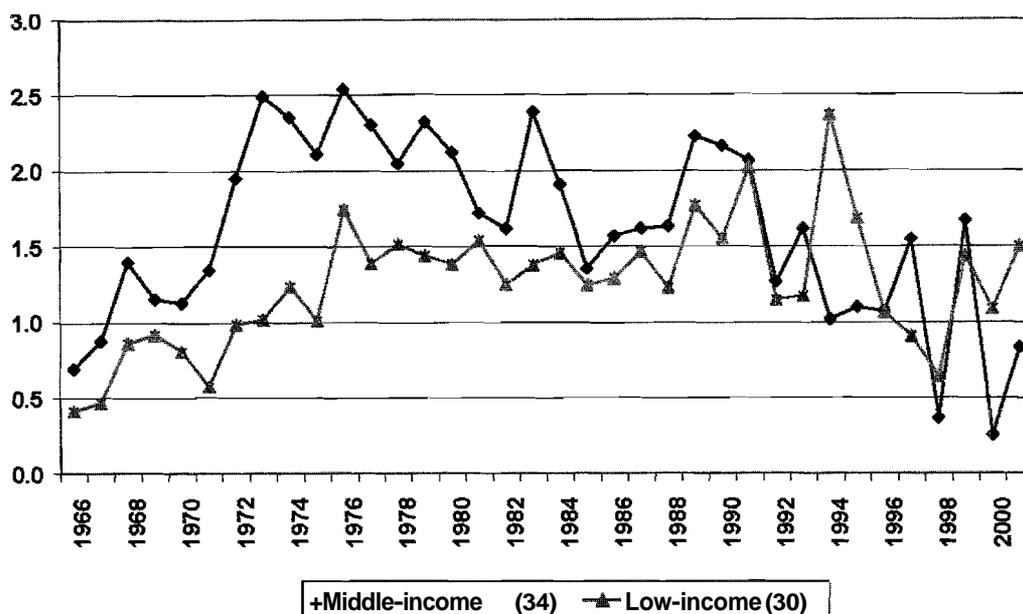
industrial countries (Montiel and Servén 2004). This high real exchange rate volatility partly reflected the high incidence of exchange rate crashes in the decade, when large devaluations were a frequent phenomenon (figure 8). Their incidence peaked in 1994, with the devaluation of the CFA franc, and in 1998, with the East Asian and Russian financial crises. Over the 1990s as a whole exchange rate crashes were slightly less frequent than in the 1980s but much more so than in the 1960s and 1970s.<sup>5</sup> High real exchange rate volatility and frequent exchange rate collapses suggest that over the 1990s progress in achieving robust nominal exchange rate arrangements was limited.

### Summary

Over the 1990s developing countries made notable progress on fiscal consolidation and limiting inflation. Improved fiscal and nominal stability helped attain a modest reduction in output volatility. These achievements were also facilitated by a somewhat more stable external environment:<sup>6</sup> the volatility of the terms of trade declined in all developing regions, in most cases to levels comparable to those of the 1960s, and capital flow volatility also fell, although to a more limited extent.

But the situation is far from rosy. In terms of outcome variables, developing countries remain much more unstable than developed ones. Moreover, extreme volatility

**Figure 7.** Developing Countries' Median Seigniorage Revenues, by Country Income Group, 1966–2001 (Percent of GDP)



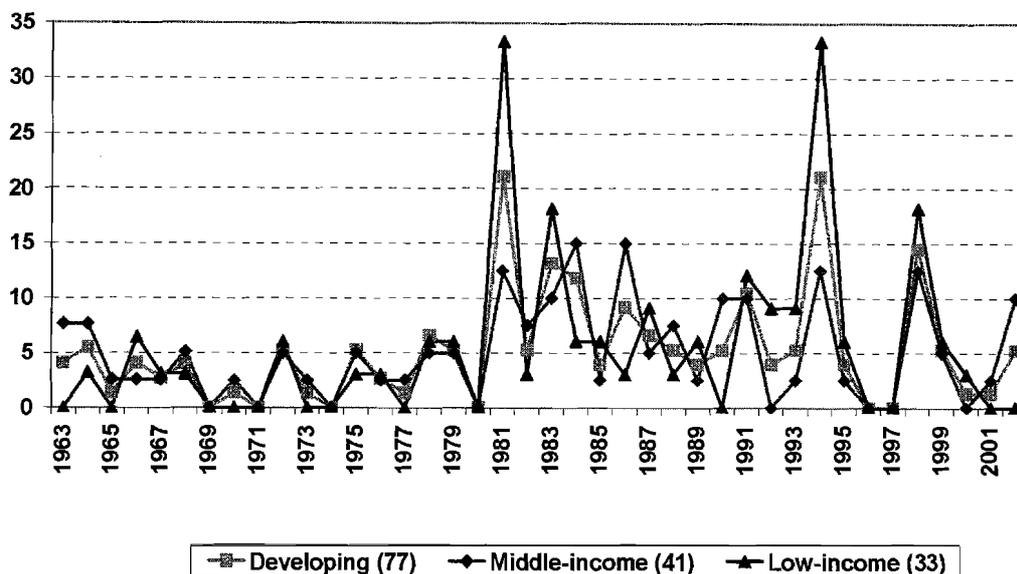
Source: IMF (various years); World Bank (various years).

accounted for a larger share of total volatility in the 1990s than before, which is consistent with evidence suggesting that currency crashes and sudden stops in capital inflows did not tend to decline during the 1990s (Montiel and Servén 2004). The situation is therefore one of dramatic policy improvements in some areas, more moderate improvements in the stability of macroeconomic outcomes, and persistent vulnerability to extreme macroeconomic events. The next section uses these findings to interpret the growth performance of developing countries during the 1990s.

## Assessing the Experience of the 1990s

The previous section has shown that macroeconomic policies and macroeconomic stability in developing countries improved along several important dimensions during the 1990s. These improvements were driven largely by the quest for higher growth. Yet as Pritchett (2004a) has argued, the growth payoff fell short of expectations. To examine why, this section briefly reviews the analytical links between macroeconomic stability and economic growth and then interprets the experience of the 1990s in the context of that analytical framework.

**Figure 8.** Share of Developing Countries Undergoing Exchange Rate Crises, 1963–2002 (Percent of Developing Countries)



*Note:* An exchange rate crisis is defined as in Frankel and Rose (1996): a depreciation of the (average) nominal exchange rate that exceeds 25 percent, exceeds the previous year's rate of nominal depreciation by at least 10 percent, and is at least three years away from any previous crisis.

*Source:* IMF (various years).

### *From Stability to Growth*

Theory suggests that the link between macroeconomic policy stability and growth has three components. First, the direct contribution that policy stability can make to growth (by ensuring that policy itself does not become an additional source of instability) likely depends on the institutional setting, because what matters for investment decisions is not only whether policy realizations are favorable today but also the perceived likelihood that appropriate policies will be repeatedly implemented in the future. To have a significant impact on growth, therefore, actual gains in macroeconomic stability need to be viewed by the private sector as indicative of a permanent change in the macroeconomic policy regime.

Second, the potential indirect contribution of policy stability to growth—by promoting the stability of macroeconomic outcomes—likely depends on the economy's degree of macroeconomic fragility, that is, the extent to which even relatively minor shocks can have large effects on the economy. On the one hand, fragility may make it too costly to deploy stabilization policies for the fear of potentially adverse effects, resulting in policy paralysis; on the other hand, fragility can mean that the instability

that policy has to counter may become so severe that feasible policy adjustments are unable to counter it.

Third, as already stressed, growth does not depend only on macroeconomic stability. The effectiveness of stability in outcomes in promoting economic growth likely depends on a variety of growth determinants, including microeconomic factors such as the definition and enforcement of property rights and the prevalence of market-oriented incentives that are jointly required for markets to perform their allocative role. The rest of this section evaluates the reform agenda of the 1990s from this three-part analytical perspective.

### *How Much Progress Was Really Achieved in the 1990s?*

As documented above, on the whole there were significant achievements in terms of stability in the traditional macroeconomic policy sense during the late 1980s and early 1990s. But these achievements were not universal, they were not always grounded on solid institutional foundations to guarantee their permanence, and they rarely translated into a more effective use of macroeconomic policies as stabilization instruments.

A useful framework within which to discuss these issues is the fiscal solvency condition,  $PV(T - G + dM) \geq B(0)$ , which requires the present value (*PV*) of primary surpluses ( $T - G$ ) and seigniorage revenue ( $dM$ ) to be at least as large as the government's outstanding stock of net debt.

From a macroeconomic policy viewpoint, stability requires the authorities to take a monetary and fiscal policy stance consistent with maintaining fiscal solvency at low inflation, while leaving some scope to mitigate the impact of real and financial shocks on macroeconomic performance. Obviously, the first requirement imposes constraints on the magnitudes of both the primary deficit and its money financing, while the second refers to the profiles of monetary and fiscal policy over the business cycle.

Most important, these requirements apply not only to current policies but also to future ones, as implied by the *PV* term in the expression. Indeed, one of the key dilemmas for macroeconomic policymaking is precisely how to ensure and convey to the private sector that future policies will abide by the requirements of solvency and low inflation without having to surrender the short-run stabilization capability of monetary and fiscal policy, that is, the tradeoff between credibility and flexibility. As discussed later, many of the achievements and disappointments of the 1990s relate to the search for lasting solutions to this dilemma.

Reassessing developments during the 1990s in light of the above expression leads to six key observations:

- A comfortable perception of fiscal solvency has yet to be established in most countries.
- Improved fiscal balances have often been achieved with stopgap measures or in ways inimical to growth and welfare.

- In many countries fiscal policy remains destabilizing.
- Lasting nominal stability remains to be credibly established.
- The transition to robust exchange rate arrangements has been anything but smooth.
- The reform agenda proved to be incomplete.

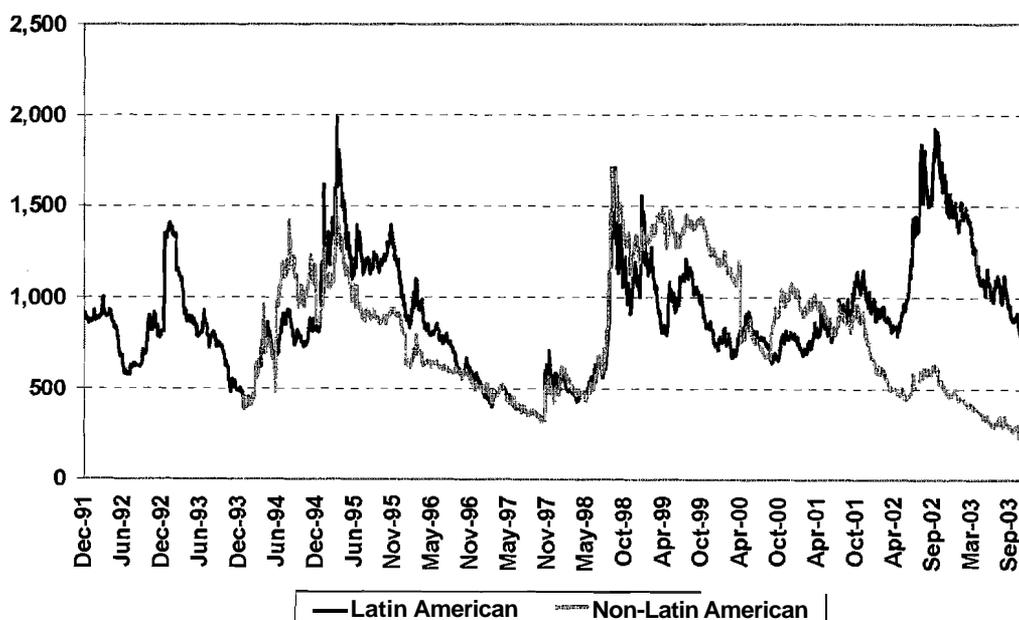
*A comfortable perception of fiscal solvency remains to be established in most countries.* Despite the trend toward lower fiscal deficits documented earlier, public debt ratios remained high in most developing countries, showing little decline during the 1990s. For developing countries with available data, the median public debt to GDP ratio remained in the 50–60 percent range over the decade. A decline through 1997 was followed by a rising pattern, so that by 2001–02 the median developing country debt ratio exceeded the 1990–2001 level—as did the median industrial country debt ratio. On the whole, for the 46 low- and middle-income countries in the sample debt ratios rose in 24 countries and fell in 22 countries.

This persistence of high debt over the 1990s and its upward drift at the end of the decade have five main causes. First, improvements in fiscal performance, as measured by reductions in primary deficits, were not universal.<sup>7</sup> Second, in many cases the pressure of weak public finances on debt accumulation was aggravated by attempts at rapid disinflation, which implied a drop in deficit monetization. Without an equally rapid correction of the primary deficit, debt issuance was left as the only available source of financing. Empirically, this is confirmed by the fact that over the 1990s disinflation shows a statistically significant association with subsequent rises in debt ratios. Third, in several countries that did achieve a fiscal adjustment over the 1990s, most public debt accumulation reflected the cost of banking system bailouts. (The realization of other contingent liabilities and the recognition of hidden ones were also significant sources of debt accumulation in some countries, such as Argentina; see Mussa 2002.) Indeed, some of the banking crises of the 1990s—especially those in East Asia in 1997—ranked at the top of the historical record in terms of fiscal impact. Fourth, where the bulk of public debt was denominated in (or indexed to) foreign currency, large real exchange rate depreciations were another major factor behind the upward trend in debt stocks in the late 1990s.<sup>8</sup> A fifth factor behind persistently high debt was the high level of real interest rates in many countries, particularly in the late 1990s, which largely reflected a lack of credibility in their stabilization efforts, documented below. Excessive reliance on short-maturity debt made some countries' overall fiscal outcomes—and thus their rates of public debt accumulation—highly sensitive to changes in domestic interest rates. Thus, in some countries (notably Brazil) high real interest rates contributed to a rapid pileup of public debt that further weakened perceptions of solvency and macro-economic stability.

In terms of the solvency constraint introduced earlier, the bottom line is that through all these channels, increases in the observed value of the primary surplus  $T - G$  were not enough to lower the public debt and establish a comfortable perception of fiscal solvency in many countries. A strong indication that solvency perceptions remained shaky in the 1990s is the fact that default risk premiums, as measured by sovereign borrowing spreads in international markets, remained highly volatile for most emerging market economies (figure 9). The evidence suggests that default risk depends not only on debt burdens but also on investors' perceptions about the quality of borrowers' policy and institutional framework (Kraay and Nehru 2003). Thus, the volatility of risk premiums likely reflected—among other factors—the market's lack of confidence in borrowers' commitment to stability.

But perceptions of high default risk are not just a symptom of perceived vulnerability; they also indirectly undermine durable outcome-based macroeconomic stability by creating macroeconomic fragility. In particular they hamper countries' ability to conduct stabilizing policy: when default risk is perceived to be high and very sensitive to changes in circumstances, attempts to run deficits at times of cyclical contraction may be viewed with suspicion and result in large jumps in risk premiums (and thus borrowing costs), discouraging the use of countercyclical fiscal policy (see Calderon, Duncan, and Schmidt-Hebbel 2003 for empirical confirmation).

**Figure 9.** Emerging Markets Bond Index for Latin American and Other Borrowers (Basis Points)



Source: JP Morgan (various years).

Moreover, the scope for independent monetary policy can also be severely constrained by the impact of changes in monetary stance on the cost of public debt through the associated changes in the nominal exchange rate and interest rate.

*Improved fiscal balances have often been achieved with stopgap measures unlikely to be sustainable or in ways inimical to growth and welfare.* Weaknesses in fiscal adjustment were not limited to the fact that increases in debt often offset improvements in primary surpluses. Often the improvements were likely to be perceived as purely temporary—because the measures behind them were transitory, because they reflected accounting transactions that had no effect on solvency, or because they directly compromised future growth and welfare. In terms of the solvency constraint above, such adjustments often had a significant impact on the current deficit but had little effect (or even an adverse one) on the path of future deficits.

In some instances, especially during the early part of the 1990s, fiscal adjustments reflected a rise in revenues from a temporary boom in tax bases—for example, a consumption boom fueled by a transitory surge in capital inflows in an economy whose tax system was dominated by the value added tax. When the boom ended abruptly, a major fiscal gap opened in the recession. There is evidence that this mechanism had a significant role in some emerging markets in the 1990s (Talvi 1997). Elsewhere a variety of accounting measures improved conventional debt indicators without making substantive progress toward fiscal solvency. Common devices included one-time asset sales to finance the retirement of public debt (which in principle implies no change in government net worth) and replacement of explicit debt with contingent liabilities (for example, granting debt guarantees rather than subsidies to public firms). Measures such as these result in improvements in a benchmark closely watched by investors and international financial institutions—gross public debt—but have no effect on solvency. In other words, they represent illusory fiscal adjustment (for example, see Easterly 1999 and Easterly and Servén 2003).

More generally, in many fiscal adjustment episodes, the focus on the quantity of adjustment was not matched by a comparable emphasis on its quality. The attention given to public spending composition and to its implications for growth and welfare has often been limited. This disinterest sometimes resulted in adjustment at the cost of basic social needs—for example, by giving inadequate protection to critical social expenditures (IMF 2003).

More often than not, productive public expenditures (on such items as human capital formation and infrastructure) are compressed in the process of fiscal adjustment, mostly because the emphasis on cash deficits and debt discourages projects whose costs are borne upfront but whose returns accrue only over time. Such projects have the same impact on the government's short-term financing needs as pure consumption or any other spending item, even though their impact on solvency is quite different because, unlike consumption, they involve creating assets

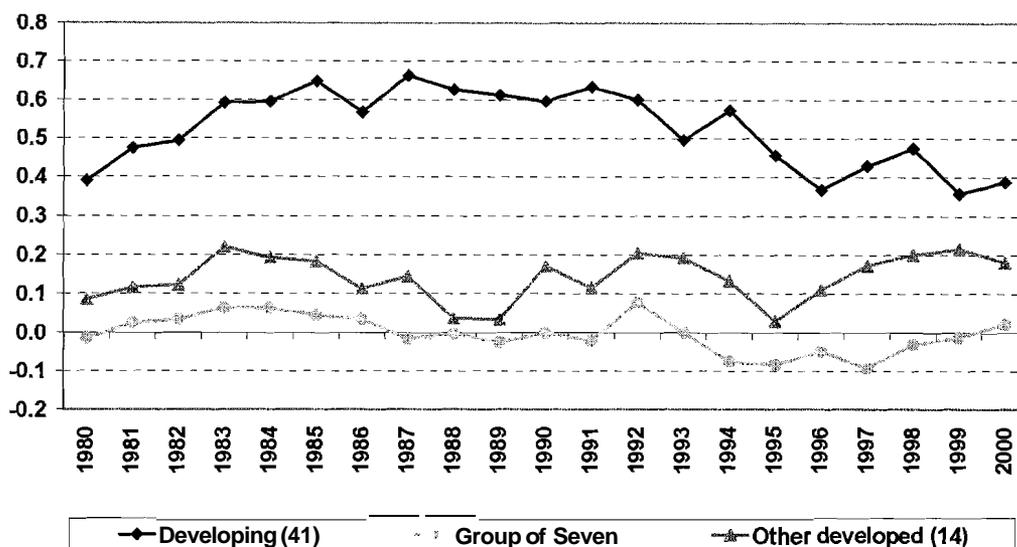
that yield future revenues—be it directly or in the form of augmented tax collection resulting from higher output levels. Conventional fiscal aggregates (such as the primary or the overall surplus) closely monitored by the international financial institutions and investors ignore this distinction, and thus fiscal adjustment tends to have an anti-investment bias amply documented in both developed and developing countries (Blanchard and Giavazzi 2003; Easterly and Servén 2003). To the extent that reduced investment lowers growth and hence future tax bases, such bias can have adverse consequences for growth—or even for fiscal solvency itself. The experience of Latin America, where declining public infrastructure spending accounted for the bulk of the fiscal correction achieved by some of the region's major countries in the 1990s, provides a good example of this perverse dynamic.

*In many countries fiscal policy remains destabilizing.* As is well known, fiscal policy in developing countries tends to be procyclical, expanding in booms and contracting in recessions. Empirical estimates show that a 1 percent increase in GDP growth tends to raise the growth rate of public consumption spending, for example, by about 0.5 percentage points in developing countries. The corresponding figure for industrial countries is much smaller (around 0.15), and for the largest of them (the Group of Seven countries), the response of public consumption is actually negative (Talvi and Vegh 2000; Lane 2003). By this measure, fiscal procyclicality in developing countries peaked in the 1980s and declined somewhat over the 1990s—but still remains much higher than in industrial countries (figure 10). Indeed, procyclical fiscal policy played a key role in some of the major crises of recent years, Argentina being a prime example (Mussa 2002; Perry and Servén 2003).

*Lasting nominal stability remains to be credibly established.* As shown earlier, developing countries substantially reduced deficit monetization during the 1990s. But whether price stability can be sustained in many of them remains to be established. As the government's intertemporal budget constraint indicates, the roots of inflation are ultimately fiscal. Thus, while a transitory reduction in deficit monetization can be achieved in a variety of ways, unless durable increases in the primary surplus are somehow institutionalized, continuing pressures on the government budget will result in debt accumulation that will in turn create pressures for monetization.

Indeed, during the 1990s, many countries' reductions in deficit monetization were not accompanied by lasting solutions to fiscal problems. In some cases (for example, Argentina, Brazil, Ecuador, Mexico, the Russian Federation, and Turkey) reduced inflation rates were achieved by stabilizing exchange rates. While improvements in price performance made reductions in money growth rates possible in these cases, sustainability remained questionable in all of them. In most cases continued fiscal pressures were accompanied by real exchange rate appreciations and increases in real interest rates, leading to a pileup of public debt and calling into

**Figure 10.** Cyclical Behavior of Public Consumption, by Country Income Group, 1980–2000



*Note:* The figure shows the median of country-specific coefficient estimates obtained by regressing the rate of growth of public consumption on the rate of GDP growth (plus a constant) over 15-year rolling windows.

*Source:* World Bank (various years).

question the sustainability of the stabilizations. In Argentina and Ecuador the lack of fiscal discipline led to the adoption of hard exchange rate pegs (a currency board in Argentina and dollarization in Ecuador) in the hope that they would somehow harden government budget constraints. Their failure to do so shows that such quick fixes are not enough to achieve lasting nominal stability without an independent commitment to responsible fiscal policies. In this way, Brazil, Mexico, and Turkey's exchange rate–based stabilizations that relied on soft pegs eventually resulted in currency crises that gave way to short bursts of accelerated inflation.

In view of this experience some countries adopted an alternative institutional arrangement during the 1990s, relying on an independent domestic central bank with a commitment to price stability. Like a fixed nominal exchange rate, such an arrangement works in principle by committing the central bank to a low value of deficit monetization ( $dM$ ), thereby imposing a hard budget constraint on the fiscal authorities and forcing them to adjust the primary deficit ( $T - G$ ) to the requirements of price stability. For such an arrangement to be effective in promoting lasting price stability, the central bank has to be committed to price stability and able to resist pressure for monetization from the fiscal side (that is, it has to avoid fiscal dominance and achieve true independence from the finance ministry). But establishing a truly independent and effective central bank has not been a straightforward matter either. The creation of independent central banks in Venezuela in 1989 and in Mexico in

1993, for example, did not prevent the substantial political pressure for credit creation that contributed to currency crises in the first half of the 1990s.

How successful have developing countries been in creating a credible commitment to nominal stability? One way to infer the private sector's expectations for nominal stability is by observing its behavior, for example, the prevalence of dollarization. Since agents can partly protect themselves from nominal instability by denominating their assets in foreign currency, improved confidence in nominal stability should reduce dollarization, even though perceptions of nominal instability are not the only factor behind financial dollarization.<sup>9</sup> However, many developing countries remained heavily dollarized by the end of the 1990s, and the median degree of dollarization of bank deposits among low- and middle-income countries actually increased over the 1990s (IMF various years, 2002; Reinhart, Rogoff, and Savastano 2003). The contrast with high-income countries is stark: their much lower degree of deposit dollarization showed little change over the same period.

Ex post real interest rates may be another indicator: they tend to be high when actual inflation falls short of expectations and when inflation uncertainty is high. Although real interest rates declined in industrial countries during the 1990s, this was not the case in developing countries, where high real interest rates persisted and were higher at the end of the decade than at the beginning (Montiel and Servén 2004).

As already noted, both dollarization ratios and ex post real interest rates reflect a variety of factors in addition to the perceptions of nominal instability, so this evidence is only suggestive. But other indicators point in the same direction. As an extreme example, the currency premium on the Argentine peso was positive throughout the 1990s and became very large at times of turbulence despite the supposedly irrevocable peg to the dollar enshrined in Argentina's Convertibility Law (Schmukler and Servén 2002).

*The transition to robust exchange rate arrangements has been anything but smooth.* Price stability refers not only to stability in the purchasing power of domestic currency over goods and services but also to an appropriate level of purchasing power over foreign exchange. However, recent progress toward robust exchange rate regimes has been uneven in developing countries. Indeed, it probably was an early casualty of the search for macroeconomic stability. As already discussed, many countries adopted exchange rate–based stabilization strategies as a supposedly quick recipe for disinflation. These not only meant the adoption of single currency pegs but also made such pegs very difficult to adjust, since the credibility of the entire stabilization program was tied up with the stability of the peg. In effect, the defense of the peg sometimes became an end in itself, even when it was evident that it had outlived its usefulness. More flexible exchange rate arrangements—that is, arrangements lacking a pre-announced peg, with or without extensive central bank intervention—have too often been adopted only in the aftermath of currency crises.

The late 1990s showed that hard exchange rate pegs—that is, dollarization and currency boards—are not a speedy shortcut to fiscal orthodoxy and nominal stability in lieu of the slow and painful buildup of credibility required when countries rely on an independent monetary policy. In particular, the Argentine episode showed the threat to stability posed by inflexible exchange rates, which made adjustment to real disturbances exceedingly difficult. These shackles eventually undermine the sustainability of such rigid arrangements. Though less well known, the experience of the CFA franc during the first half of the 1990s is another example of this situation.

*The reform agenda proved to be incomplete.* The preceding observations suggest that, as far as fiscal solvency and price stability are concerned, the reform agenda of the 1990s left much to be desired. But the agenda was also deficient in its very design, because the macroeconomic stability that fiscal solvency and price stability are supposed to deliver was undermined by leaving in place—or worse yet, creating—important sources of macroeconomic fragility.

A particular area of fragility in which the policy-based stability agenda was incomplete is financial sector soundness. While research shows that an efficient domestic financial system is important for growth, the experience of the 1990s strongly suggests that a sound one is indispensable for macroeconomic stability. The macroeconomic reform agenda of the early 1990s was incomplete in that the central role of the financial system for macroeconomic stability was often ignored—even though it should have been clear in light of the Southern Cone crises of the early 1980s. Thus to the standard policy-oriented prescriptions for stability—a solvent fiscal stance, low and stable money growth, and robust exchange rate policies that nevertheless allow adjustment to shocks—it is necessary to add policies that foster a sound financial system. Indeed, in the wake of the crises of the 1990s the IMF redefined its core competencies to include fiscal, monetary, exchange rate, and financial sector policies.

Stability in this particular sense—that is, ensuring a sound domestic financial system—was clearly not widely achieved by developing countries during the 1990s. As a result, an important source of macroeconomic fragility was not only left in place but may have even been magnified, for reasons to be explained. Inadequate attention to financial sector soundness often resulted in a domestic economic environment in which institutional problems involving moral hazard were rife, rendering both public and private balance sheets highly vulnerable to changes in the environment (interest rate and exchange rate changes) and posing a major obstacle to outcome-based stability in several major countries.<sup>10</sup> The proliferation of financial crises in the 1990s reflects in part this missing piece of the reform agenda. Indeed, the incidence of systemic banking crises was even higher in the 1990s than in the 1980s, particularly in the second half of the decade (Bordo and others 2001).

But the frequency and the severity of crises were also affected by an important change in the economic environment—namely, increased capital mobility. This was

another key source of fragility, making economies vulnerable to sudden shifts in capital flows. In fact, the combination of unsound policies in the financial sector and open capital accounts helps explain many characteristics of the crises of the 1990s.

First, many of these crises were twin crises, simultaneously involving currency and banking collapses, often characterized by banking problems preceding a currency crash, which then fed back into a full-blown financial crisis (Kaminsky and Reinhart 1999). There is evidence that twin crises are usually much more costly in terms of output than standard banking-only or currency-only crises (Bordo and others 2001). Second, many of these crises proved hard to foresee on the basis of standard macroeconomic imbalances. The hardest—especially the Mexican and East Asian crises—occurred where the main vulnerabilities concerned financial, rather than macroeconomic, variables and took the form of balance of payments runs similar to traditional bank runs." Third, many of these crises were surprisingly severe. The deepest ones involved serious problems in the financial sector (East Asia, Ecuador, Mexico, and Turkey), in private sector balance sheets (Argentina and East Asia), and with fiscal insolvency (Argentina and Ecuador). Where none of these problems was present and events took the form of a simple currency crash (Brazil), crisis-induced economic contraction was not as severe.

### *The Growth Payoff*

While the improvements in macroeconomic policies were limited—as the preceding discussion has shown—growth rates have indeed risen relative to the 1980s in many developing countries. The achievement is only a modest one, however, since growth in the 1980s was generally low, and for most countries growth rates over the 1990s remained well below those over the 1960s and 1970s. Indeed, of the 77 developing countries with complete data, only 28 achieved growth rates over the 1990s that exceeded those over the 1970s and only 24 achieved growth rates that exceeded those over the 1960s. But is this growth payoff commensurate with the progress on macroeconomic stability, or is it disappointing? There are several reasons to believe that the growth payoff was indeed commensurate with what was actually achieved by reform.

First, as argued above, the growth payoff from macroeconomic stability depends on its perceived permanence. But, as discussed in the previous section, often progress on macroeconomic stability was based on policy changes that were not perceived to be durable or that failed to reform the institutions making macroeconomic policy. In this sense, the growth payoff expected from the stability that was actually achieved may have been overstated. Moreover, a vicious circle may have taken hold in some countries, with the social consensus that made the policies possible—and that is needed to make them sustainable—faltering in the absence of a fairly prompt growth payoff.

Second, the search for macroeconomic stability—narrowly defined—may in some cases have actually been inimical to growth. As already noted, a preoccupation with

reducing inflation induced some countries to adopt exchange rate regimes that ultimately conflicted with outcome-based stability. In other cases, as shown previously, a single-minded pursuit of macroeconomic stability may have come at the expense of growth-enhancing policies (for example, an adequate provision of public goods) and social investments that might have both increased the growth payoff and made stability more durable. From this perspective, some economies may well have been over-stabilized in both microeconomic and macroeconomic senses. From a microeconomic perspective the presumed stability gains from further fiscal adjustments may not have justified the costs of forgoing key social and productive expenditures. From a macroeconomic perspective the narrow focus on stability may have precluded more progress toward countercyclical policies. The contrast between the significant fiscal adjustment achieved by most developing countries and the persistence of outcome-based instability suggests that this factor may have been important.

Third, aside from whether the search for macroeconomic stability worked at cross-purposes with the search for higher growth, the incomplete macroeconomic reform agenda failed to bring about the reduction in macroeconomic fragility required to fully translate policy-based stability into outcome-based stability. Although overall macroeconomic volatility decreased among developing countries, extreme volatility actually rose during the 1990s, reflecting largely a spate of crises during the decade. Moreover, the adverse impacts of extreme volatility on growth appear to exceed those of normal volatility.<sup>12</sup> Thus, the growth payoff of the macroeconomic policy improvements achieved in the 1990s was limited not only by their weak institutional underpinnings but also by the extreme outcome-based instability that emerged during the decade mainly as a result of the fragilities overlooked by an incomplete reform agenda.

Fourth, as argued earlier, while macroeconomic stability may facilitate growth when other forces are driving the growth momentum—that is, when macroeconomic instability is the binding constraint on growth—macroeconomic stability is not enough to drive the growth process when other essential ingredients are lacking—that is, when other constraints are binding. These constraints involve the various policies and institutions that shape the opportunities and incentives to engage in growth-enhancing activities (Pritchett 2004b). They include secure property rights and market-oriented microeconomic incentives as well as proactive government interventions to overcome informational externalities and coordination failures (Rodrik 2004). The importance of these complementary factors may not have been sufficiently appreciated early in the decade.

In sum, there is little reason to expect a simple direct association between macroeconomic stability and growth, even if stability as measured by commonly used macroeconomic policy indicators is achieved. From this perspective, the limited growth payoff that emerged from the gains in macroeconomic stability achieved during the 1990s may not be all that surprising.

## Summary: Lessons from the 1990s

This article does not take a position on why the implementation of reform in developing countries during the 1990s proved to be flawed in the ways emphasized. Perhaps, as others have suggested, the problem was the absence of a clear order for reform priorities in a context where developing countries have limited political and administrative capacity for reform (Hausmann, Rodrik, and Velasco 2005). This may account for many countries' inability to put in place the institutional underpinnings of macroeconomic stability in cases where macroeconomic instability was indeed the binding constraint on growth, and for other countries' failure to implement microeconomic reforms with a larger potential growth payoff when macroeconomic instability was not the binding constraint. This section instead draws some lessons from the reform experience as implemented.

This experience offers several lessons for the future. A central one is that the old verities concerning the importance of macroeconomic stability still hold true. While macroeconomic policy realizations are not all that matter for promoting economic growth, they clearly do matter. Perceived fiscal insolvency, high and unstable inflation, and severely overvalued real exchange rates remain reliable ingredients for extreme instability and slow growth. But the 1990s also showed that in addition to macroeconomic policy realizations, three other ingredients are critical for growth: (a) the institutional framework within which fiscal, monetary, and exchange rate policies are formulated; (b) the degree to which macroeconomic fragilities are avoided; and (c) the extent to which complementary pro-growth reforms are implemented. These elements are reviewed briefly below.

### *Institutions for Macroeconomic Policy Formulation*

The institutional context in which traditional macroeconomic policies are formulated is critical when resolving the tradeoff between policy credibility and flexibility. Both are required for the durable outcome-based stability that ultimately matters for economic growth.

Procyclical fiscal policies arise in developing countries because without strong budgetary institutions a "tragedy of the commons" phenomenon sets in during prosperous times, when government revenues are high: because no claimant on the government's budgetary resources has an incentive to internalize the need for fiscal solvency, political imperatives cause the government to spend all its resources (and even to borrow) during booms, leaving little margin of solvency to draw on to finance fiscal deficits when times are bad. A mechanism that makes it politically possible to ensure prudent fiscal responses to favorable shocks is required in this context.

The specific mechanism best suited to the job depends on country circumstances. It may require, for example, reforming budgetary institutions to centralize budgetary

authority in the finance ministry rather than in line ministries or the parliament or implementing fiscal rules that force claimants to government resources to respect the government's intertemporal budget constraint. These transparent fiscal rules may be embodied in the country's constitution or subject to change only by legislative supermajorities, with penalties stipulated for noncompliance (Perry 2003). Alternative proposals have focused on independent fiscal policy councils to set annual deficit limits, modeled along the lines of independent central banks. However, such institutional arrangements need to balance credibility and flexibility. Simpler rules may be more transparent and thus more easily verifiable, but they need to be designed to allow sufficient flexibility for fiscal policy to react to a changing economic environment. Overly rigid rules are unlikely to be sustainable or credible—as shown by the increasing pressures to revise the European Stability Pact because of its neglect of the macroeconomic cycle.

With respect to monetary policy and exchange rate regimes, the evidence indicates that low and stable inflation is conducive to economic growth, and theory suggests that it is most important in this regard for the private sector to be convinced that low and stable inflation is a permanent feature of the economic environment. As in the case of fiscal credibility, an appropriate institutional underpinning for price stability is required to generate such a perception. However, as shown earlier, a key lesson of the 1990s was that purely monetary arrangements cannot achieve such a perception, because they are not sufficient to discipline fiscal policy. Fiscal credibility is a necessary condition for monetary credibility, and not even the most rigid monetary arrangements (a currency board or *de jure* dollarization) guarantee hard government budget constraints. In short, no institutional shortcuts of a purely monetary nature can achieve credible price stability. Instead, the task is to choose monetary arrangements that can best complement reformed fiscal institutions in achieving a desirable tradeoff between credibility and flexibility.

Again, the optimal institutional responses to this challenge are likely to be country specific. The experience of the 1990s, however, suggests that a monetary arrangement often suitable for this task features an independent central bank that operates a floating exchange rate and commits to a publicly announced inflation target. This arrangement has the important advantages of flexibility (since the central bank is not constrained in how it attains its inflation target) and commitment (since the central bank's prestige is publicly put on the line by announcing such a target). Most important, floating exchange rates and inflation targets allow the domestic authorities to establish anti-inflationary credibility the hard way—that is, by establishing a track record—rather than by attempting to import it through some form of exchange rate peg.<sup>13</sup>

### *Robustness: The Scope of the Macroeconomic Reform Agenda*

The proliferation of crises during the 1990s has made it clear that beyond an appropriate institutional setting for formulating fiscal, monetary, and exchange rate

policies the reform process in developing countries also needs to attend to robustness issues to achieve the stable macroeconomic environment sought through fiscal and monetary reforms. This includes, in particular, policies directed toward the domestic financial system and capital account, which have been shown to have important implications for macroeconomic fragility and thus for outcome-based macroeconomic stability.

As has been widely recognized, an appropriate institutional framework in the domestic financial sector involves (a) clear and secure property rights; (b) an accessible, efficient, and impartial legal system to enforce contracts; (c) appropriate legal protection for creditors; (d) well-specified accounting and disclosure standards; (e) a regulatory system that screens entrants while encouraging competition, imposing adequate capital requirements, and preventing excessively risky lending; and (f) a supervisory system that can effectively monitor the lending practices of domestic financial institutions. The key lesson is that the pace of liberalization for domestic financial systems that have not already been liberalized should be modulated to reflect the quality of the institutional framework governing the domestic financial sector and that improving the quality of this framework deserves high priority in the macroeconomic reform agenda.

In contrast to the consensus that has emerged on domestic financial reform, managing a country's integration into international financial markets remains a controversial part of the institutional agenda. Despite strong theoretical arguments concerning the gains from capital account openness, the empirical evidence on whether it has in fact been conducive to enhanced growth and reduced consumption volatility is inconclusive (on growth, see Edison and others 2002; on consumption volatility, see Kose, Prasad, and Terrones 2003). The desire to avoid macroeconomic fragility makes a strong case for institutional arrangements for the capital account that at least prevents maturity mismatches in a country's external balance sheet, since such mismatches can leave the country vulnerable to creditor runs analogous to bank runs. These runs played a key role in the East Asian crisis (Rodrik and Velasco 1999). The question is how to preclude them. One difficulty is that short maturities are attractive to creditors as a means of monitoring borrowers and controlling their behavior precisely when asymmetric information and moral hazard problems are serious. Under these circumstances, therefore, short-maturity borrowing will arise endogenously because it will be substantially less costly to borrowers than long-term loans. Mismatches may reflect not only an inadequate borrowing strategy but also the reluctance of investors to lend long-term in the face of a macroeconomic financial framework deemed suspect. The problem is, of course, that voluntary short-maturity loans between private parties fail to take into account the social costs associated with the risk of creditor runs.

Again, there are alternative ways of meeting this challenge. One way to deal with maturity mismatches is for the public sector to accumulate large stocks of foreign

currency (foreign exchange reserves) to offset liquid liabilities incurred by the private sector. This approach is being pursued by many countries, but it is expensive. Holding large volumes of low-yielding, short-term assets instead of (illiquid) long-term investment entails serious opportunity costs and leaves in place incentives that give rise to short-term borrowing. Under this strategy the costs of insuring against creditor runs are ultimately borne by taxpayers. An alternative is to discourage the private sector from incurring short-term external liabilities in the first place through restrictions on short-term capital inflows or to make those liabilities effectively less liquid at times of crisis through restrictions on short-term capital outflows. Because both of these policies tend to increase the cost of short-term loans, they effectively operate by internalizing the systemic costs associated with the risk of creditor runs. The available evidence suggests that inflow restrictions such as unremunerated reserve requirements (for example, the Chilean *encaje*) tend to have no significant effect on the overall volume of inflows but do affect their composition, reducing the share of short-term flows in the total. This is because a uniform reserve requirement is more onerous for short-term transactions than for others (see Montiel and Reinhart 1999 for review of the cross-country evidence on the effectiveness of inflow restrictions). By contrast, evidence on the effects of restrictions on outflows is much less conclusive. Overall, it appears that restrictions on short-term capital inflows may have a role in achieving outcome-based macroeconomic stability in developing countries.

In addition to maturity mismatches, external borrowing may also create fragility in the form of currency mismatches. Developing countries may suffer from "original sin"—an inability to denominate external borrowing in their own currencies—which causes exchange rate risk to be borne by domestic residents rather than by foreign residents (Eichengreen and Hausmann 1999). In the short run the key is to promote the efficient distribution of this exchange rate risk within the domestic economy by ensuring—through regulatory means—that it is appropriately priced and therefore borne by those best able to bear it (typically agents holding foreign currency assets—including exporters—or those with a high degree of risk tolerance). In the longer term a larger role in ameliorating the problem of currency mismatches would be assumed by institutional changes that promote credible nominal stability, thus mitigating exchange rate risk. This perspective is consistent with the experience of a few emerging economies—such as Chile, Mexico, Poland, and South Africa—for which domestic currency-denominated external borrowing is becoming a reality, even if only marginally at first.

### Complementarities Among Pro-Growth Policies

But most of the burden of jumpstarting growth in developing countries must fall on pro-growth policies outside the macroeconomic arena. Aside from the fundamental

policies, securing property rights, and establishing market-oriented incentives, such policies may include, for example, an open international trade regime, national innovation policies, well-functioning factor markets, and an investor-friendly legal and regulatory environment. In some cases these policies actually facilitate reforms aimed at macroeconomic stability—for example, disinflation or the correction of a real misalignment are easier and less costly to achieve with well-functioning labor and financial markets. The key lesson is that policies of this type are mutually complementary with policies that focus on creating and preserving macroeconomic stability. An unstable macroeconomic environment tends to undermine the growth benefits of such policies. Nonetheless, the 1990s showed that macroeconomic stability alone is not enough: policies outside the macroeconomic arena are indispensable for harvesting the fruits of macroeconomic stability in the form of sustained high rates of economic growth.

## Notes

Peter Montiel is a Professor of Economics at Williams College; his email address is [peter.j.montiel@williams.edu](mailto:peter.j.montiel@williams.edu). Luis Servén is Research Manager for Macroeconomics and Growth in the Development Research Group of the World Bank; his email address is [lservern@worldbank.org](mailto:lservern@worldbank.org). This article was originally prepared as a background paper for the volume "Lessons from the 1990s." The authors are indebted to Roberto Zaghera for numerous comments and discussions throughout the preparation of the article. They also thank, without implication, Indermit Gill, Jim Hanson, Ricardo Hausmann, Ravi Kanbur, Phil Keefer, Guillermo Perry, Lant Pritchett, S. Ramachandran, and three anonymous reviewers for valuable comments on previous versions. Patricia Macchi provided excellent research assistance.

1. Easterly (2001) also viewed the crises of the 1990s as a symptom of, rather than an explanation for, the slow growth of the 1990s.

2. In recent years, there has been renewed interest, sparked by Ramey and Ramey (1995), in the adverse effects that real and nominal instability can have on economic growth, as documented by a growing empirical literature on the subject. For a recent evaluation, see Hnatkovska and Loayza (2004).

3. The decline in developing country volatility over the 1990s is also documented by De Ferranti and others (2000), Rodrik (2001), and Hnatkovska and Loayza (2004). The same result holds if volatility is measured by a robust statistic such as the interquartile range instead of the standard deviation. The decline in volatility was statistically significant: formal tests strongly reject the hypothesis that the cross-country distribution of growth volatility did not change between the 1980s and 1990s and the hypothesis that the changes in volatility across the two decades are centered at zero.

4. The decline in aggregate volatility also extends to other variables more directly related to individuals' welfare, such as income and consumption growth, although to varying extents. In particular, the volatility of private consumption growth declined relative to the previous decade but mainly in low-income countries; in middle-income countries consumption volatility remained virtually unchanged from the record highs of the 1980s.

5. In a smaller country sample (whose time coverage ends in 1997), Bordo and others (2001) also found that the frequency of currency crashes declined in the 1990s compared with the preceding 15 years.

6. The evidence for increased stability of the external environment is presented in the working paper version of this article (Montiel and Servén 2004), which also includes additional figures.

7. In India, for example, continuing large primary deficits (averaging close to 4 percent of GDP in the late 1990s) were the main factor behind persistently high debt ratios. More dramatically, fiscal vulnerabilities had a prominent role in some of the major recent financial crises: the Russian Federation in 1998, Ecuador in 1999, and Argentina in 2002. In Argentina, the expansionary fiscal stance taken during the 1995–97 boom left authorities virtually no room to adjust to the global real and financial slowdown after the Russian crisis of 1998 and to the real appreciation of the peso under the hard dollar peg (Perry and Semen 2003).

8. In Argentina and Uruguay, for example, the 2002 exchange rate collapse more than doubled the debt to GDP ratio—from 50 percent of GDP to more than 140 percent in Argentina and from 40 percent to more than 80 percent in Uruguay. Across emerging markets debt dollarization remained pervasive: the median ratio of foreign currency debt to total public debt rose over the late 1990s to exceed 55 percent by 2001.

9. The degree of real dollarization and the perceived stability of the real exchange rate also matter, as do financial system regulations and the availability of other assets that shelter investors from nominal instability, such as instruments indexed to domestic inflation, as in Chile, or short-term interest rates, as in Brazil (IMF 2002; De la Torre and Schmukler 2003). Thus the interpretation in the text should be taken as suggestive rather than as conclusive.

10. Ironically, under these circumstances incipient progress along conventional dimensions of macroeconomic stability, such as disinflation, may even have made financial crises more likely. For example, the use of the exchange rate as a nominal anchor may have encouraged agents to ignore exchange rate risk and in the case of hard pegs, as in Argentina, may have made it more difficult for regulators to induce financial institutions to factor such risk into their portfolio allocations without raising fears of a possible abandonment of the peg.

11. The recent analytical literature on crises continues to stress weak fundamentals as a prerequisite for the occurrence of crises but emphasizes the key role of ingredients such as self-fulfilling expectations and multiple equilibria in triggering them. These views assign an increasingly important role to financial system imperfections in full-blown balance of payments crises (Krugman 1999).

12. There are good reasons why crisis volatility (due to large adverse shocks) should entail greater growth wsts than normal volatility. On the one hand, with a given set of risk management mechanisms, large shocks may be more difficult to absorb than small ones. These threshold effects of volatility have been found to be empirically relevant to investment (Servén 2003). On the other hand, asymmetries built into the economy mean that negative shocks have qualitatively different consequences than positive ones. A clear example is buffer stocks—for example, bank liquidity or international reserves. Large adverse shocks (or a succession of small negative ones) can exhaust them and trigger an adjustment mechanism very different from the one associated with positive disturbances. The same applies to firms' net worth: once it becomes negative, adjustment proceeds through bankruptcies, with the corresponding destruction of productive assets. Hnatkovska and Loayza (2004) provided empirical evidence that crisis-type volatility is significantly more adverse for growth than normal volatility.

13. Such arrangements are currently maintained by Brazil, Chile, Colombia, the Republic of Korea, Mexico, Peru, South Africa, and Thailand. The longest running of these arrangements, in Chile, was remarkably successful in maintaining price stability throughout the 1990s, while avoiding severe episodes of real exchange rate volatility. More recent converts to this type of nominal institutional arrangement have also been quite successful since its (admittedly recent) adoption.

## References

Blanchard, O., and F. Giavazzi. 2003. "Improving the Stability and Growth Pact Through Proper Accounting of Public Investment." Discussion Paper 4220. Centre for Economic and Policy Research, London.

- Bordo, M., B. Eichengreen, D. Kligebiel, and M. S. Martinez-Peria. 2001. "Is the Crisis Problem Growing More Severe?" *Economic Policy: A European Forum* 16(32):51–82.
- Calderón, C., R. Duncan, and K. Schmidt-Hebbel. 2003. "The Role of Credibility in the Cyclical Properties of Macroeconomic Policies in Emerging Economies." Working Papers Central Bank of Chile, Central Bank of Chile, Santiago.
- De Ferranti, D., G. Perry, I. Gill, and L. Servén. 2000. *Securing Our Future*. Washington, D.C.: World Bank.
- De la Torre, A., and S. Schmukler. 2003. *Coping with Risk Through Mismatches: Domestic and International Financial Contracts for Emerging Economies*. Washington, D.C.: World Bank.
- Easterly, W. 1999. "When Is Fiscal Adjustment an Illusion?" *Economic Policy* 14(28):55–76.
- . 2001. "The Lost Decade: Developing Countries' Stagnation in Spite of Policy Reform." *Journal of Economic Growth* 6(2):135–57.
- Easterly, W., and L. Servén, eds. 2003. *The Limits of Stabilization: Infrastructure, Public Deficits, and Growth in Latin America*. Palo Alto, Calif.: Stanford University Press.
- Easterly, W., R. Islam, and J. Stiglitz. 2001. "Volatility and Macroeconomic Paradigms for Rich and Poor." In J. Dréze, ed., *Advances in Macroeconomic Theory*. New York: Palgrave.
- Edison, H., R. Levine, L. Ricci, and T. Sloken. 2002. "International Financial Integration and Economic Growth." Working Paper 9164. National Bureau of Economic Research, Cambridge, Mass.
- Eichengreen, Barry, and Ricardo Hausmann. 1999. "Exchange Rates and Financial Fragility." Working Paper 7418. National Bureau of Economic Research, Cambridge, Mass.
- Fitch Ratings. Various years. *Sovereign Comparator*. New York.
- Frankel, Jeffrey A., and Andrew K. Rose. 1996. "Currency Crises in Emerging Markets: Empirical Indicators." Working Paper 5437. National Bureau of Economic Research, Cambridge, Mass.
- Hausmann, R., D. Rodrik, and A. Velasco. 2005. "Growth Diagnostics." Harvard University, John F. Kennedy School of Government, Cambridge, Mass.
- Hnatkovska, V., and N. Loayza. 2004. "Volatility and Growth." Policy Research Working Paper 3184. World Bank, Washington, D.C.
- IADB (Inter-American Development Bank). 1995. *Overcoming Volatility in Latin America*. Washington, D.C.
- IMF (International Monetary Fund). 2002. *Financial Stability in Dollarized Economies*. Washington, D.C.
- . 2003. *Fiscal Adjustment in IMF-Supported Programs*. Washington, D.C.
- . Various years. *International Financial Statistics Database*. Washington, D.C.
- JP Morgan. Various years. *Emerging Market Bond Index (EMBI)*. New York.
- Kaminsky, G., and C. Reinhart. 1999. "The Twin Crises: the Causes of Banking and Balance of Payments Crises." *American Economic Review* 89(3):473–500.
- Kose, A., E. Prasad, and M. Terrones. 2003. "Financial Integration and Macroeconomic Volatility." IMF Staff Paper 03/50. International Monetary Fund, Washington, D.C.
- Kraay, A., and V. Nehru. 2003. "When Is External Debt Sustainable?" Policy Research Working Paper 3200. World Bank, Washington, D.C.
- Krugman, P. 1999. "Balance Sheets, the Transfer Problem and Financial Crises." In P. Isard, A. Razin, and A. Rose, eds., *International Finance and Financial Crises*. London: Kluwer.
- Lane, P. 2003. "The Cyclical Behavior of Fiscal Policy: Evidence from the OECD." *Journal of Public Economics* 87(12): 2661–75.
- Montiel, P., and C. Reinhart. 1999. "Do Capital Controls and Macroeconomic Policies Influence the Volume and Composition of Capital Flows? Evidence from the 1990s." *Journal of International Money and Finance* 18(4):619–35.

- Montiel, P., and L. ServCn. 2004. "Macroeconomic Stability in Developing Countries: How Much Is Enough?" Policy Research Working Paper 3456. World Bank, Washington, D.C.
- Mussa, M. 2002. *Argentina and the Fund: From Triumph to Tragedy*. Washington, D.C.: Institute for International Economics.
- Perry, G. 2003. "Can Fiscal Rules Help Reduce Macroeconomic Volatility?" Policy Research Working Paper 3080. World Bank, Washington, D.C.
- Perry, G., and L. Servén. 2003. "The Anatomy of a Multiple Crisis: Why Was Argentina Special and What Can We Learn from It?" Policy Research Working Paper 3081. World Bank, Washington, D.C.
- Pritchett, L. 2004a. "The Grist and the Mill for the Lessons of the 1990s." In G. Nanltani and R. Zagha, eds., *The Growth Experience: What Have We Learned From the 1990s* Washington, D.C.: World Bank.
- . 2004b. "Policy Reforms and Growth Performance: What Have We Learned?" In G. Nankani and R. Zagha, eds., *The Growth Experience: What Have We Learned From the 1990s?* Washington, D.C.: World Bank.
- Ramey, G., and V. Ramey. 1995. "Cross-Country Evidence on the Link between Volatility and Growth." *American Economic Review* 85(5):1138–51.
- Reinhart, C., K. Rogoff, and M. Savastano. 2003. "Addicted to Dollars." Working Paper 10015. National Bureau of Economic Research, Cambridge, Mass.
- Rodrik, D. 2001. "Why Is There So Much Economic Insecurity in Latin America?" *CEPAL Review* 73:7–30.
- . 2004. "Growth Strategies." Harvard University, John F. Kennedy School of Government, Cambridge, Mass.
- Rodrik, D., and A. Velasco. 1999. "Short-Term Capital Flows." Working Paper 7364. National Bureau of Economic Research, Cambridge, Mass.
- Schmukler, S., and L. ServCn. 2002. "Pricing Currency Risk under Currency Boards." *Journal of Development Economics* 69(2):367–91.
- Servén, L. 2003. "Real Exchange Rate Uncertainty and Private Investment in LDCs." *Review of Economics and Statistics* 85(1):212–18.
- Shvets, O. 2004. "Something Special about the 1990s?" In G. Nanltani and R. Zagha, eds., *The Growth Experience: What Have We Learned From the 1990s?* Washington, D.C.: World Bank.
- Talvi, E. 1997. "Exchange-Rate Based Stabilization with Endogenous Fiscal Response." *Journal of Development Economics* 54(1):59–75.
- Talvi, E., and C. Vegh. 2000. "Tax Base Variability and Pro-Cyclical Fiscal Policy." Working Paper 7499. National Bureau of Economic Research, Cambridge, Mass.
- World Bank. Various years. *World Development Indicators Database*. Washington, D.C.

---

# Banking and Regulation in Emerging Markets: The Role of External Discipline

---

Xavier Vives

---

*This article reviews the main issues of regulating and supervising banks in emerging markets with a view toward evaluating the long-run options. Particular attention is paid to Latin America and East Asia. These economies face a severe policy commitment problem that leads to excessive bailouts and potential devaluation of claims of foreign investors. This exacerbates moral hazard and makes a case for importing external discipline (for example, acquiring foreign short-term debt). However, external discipline may come at the cost of excessive liquidation of entrepreneurial projects. The article reviews the tradeoffs imposed by external discipline and examines various proposed arrangements, such as narrow banking, foreign banks and foreign regulation, and the potential role for an international agency or international lender of last resort.*

Liberalization and integration of financial markets have been associated with an increase in capital movements and with the financial crises. In particular, surges in foreign short-term debt have been blamed for crisis episodes in emerging economies in Asia (Thailand, Indonesia, and the Republic of Korea) and Latin America (Mexico, Brazil, Ecuador, and Argentina), as well as in the periphery of Europe (Turkey). These crises have proved costly in terms of output.

Several policy responses have been suggested. Among them have been the reduction of short-term debt, the development of stock markets, the improved regulation and supervision of domestic financial system, enhanced transparency requirements and market discipline, and the establishment of an international lender of last resort. A catalog of "solutions" has been proposed to take care of the problems of banking in emerging market economies including moving to a narrow bank system, building a currency union, and leaving banking in the hands of foreign banks and offshore institutions.

This article identifies policy responses tailored to the needs of emerging market and developing economies. The question is whether the regulatory policies and

practices of developed economies can be recommended essentially without change or whether a different policy mix is needed. A basic theme is that more acute asymmetric information problems and a weak institutional structure in emerging market economies call for policy prescriptions that differ not only from those of developed economies but also across emerging market economies.

Attention is focused on a particular consequence of the weak institutional structure in emerging countries: the lack of capacity for policy commitment. This lack of capacity for commitment may be due to the short horizons of public officials in the face of, for example, political instability. The outcome is that the government of an emerging market economy may bail out the private sector, encouraging excessive risk taking, or devalue the claims of foreign investors, discouraging their investment in the first place. Indeed, a major problem in emerging markets is the implicit or explicit guarantee of a bailout in the event of a banking crisis, as experiences in Argentina, Mexico, and Thailand show, or the use of inflation to devalue domestic currency-denominated claims.<sup>1</sup>

The result is that domestic regulation may not be enough in countries that face a commitment problem, and those countries may have to import discipline. However, some ways of importing discipline, such as increasing the role of short-term foreign debt, have costs. The article examines the tradeoffs imposed by different ways of importing discipline and classifies countries according to the desirability of doing so. It analyzes the catalog of solutions to the problems of financial systems in emerging market economies and the potential role of an international agency such as the International Monetary Fund (IMF).

## Banking in Emerging Market Economies

What makes banking regulation different in emerging market economies? Why do these countries require different regulatory and supervisory arrangements?

### *The Role of Banks and Fragility*

Banks provide transaction and payment system services, insurance, and risk sharing (transforming illiquid assets into liquid liabilities). A central function of banks is the financing and monitoring of entrepreneurial projects, which are illiquid and opaque because of asymmetric information problems such as moral hazard and adverse selection.<sup>2</sup> Some entrepreneurial projects cannot obtain market financing because no credible information on them can be conveyed to the public domain. A bank can accumulate relationship-specific skills to monitor those projects and be able to finance them. In this way the banking system helps overcome problems associated with asymmetric information in an economy.

Asymmetric information problems are bound to be more acute in emerging market and developing economies. The production of information is more problematic in emerging market economies because of institutional factors. Indeed, emerging market economies fare poorly on the indicators of rule of law, the protection of property rights, and accounting standards, pointing to aggravated moral hazard and adverse selection problems. Furthermore, the production of information, which typically involves a fixed cost, is discouraged by the normally small size of the emerging market.

One direct consequence of the enhanced asymmetric information problems is that the financial system is less developed because the cost of setting up well-functioning markets is higher. Arms-length finance just does not work. A derived consequence is that the role of the banking system, in particular the monitoring of entrepreneurial projects, is much more crucial. Indeed, for most companies in an emerging market economy the only possible source of finance, except for earnings, is bank loans. Banks and their monitoring capacity are therefore at the center of economic development, and their potential fragility may dramatically worsen downturns. The crises of Mexico, East Asia, and Russia provide examples, as well as the more recent crisis in Argentina.

Why are banks fragile? The essence of banks is that they create liquidity, which leaves them vulnerable to runs. Banks protect entrepreneurs that need finance from the liquidity needs of depositors and investors. There are different versions of the story, but this is the cornerstone of modern banking theory (Diamond and Dybvig 1983; Holmstrom and Tirole 1997, 1998; Diamond and Rajan 2001).<sup>3</sup> Firms may be unable to obtain funding because of asymmetric information, as they do not have enough pledgeable income (the fraction of their return that can be committed to be paid to outsiders). Banks come to the rescue, for example, by creating liquidity holding collateral and committing to make payments (Holmstrom and Tirole 1997, 1998). In short, the standard deposit contract and loan provision to opaque entrepreneurial projects are complementary and central to the function of a bank.

Short-term debt—a deposit redeemable at par—leaves banks exposed to failure when returns are low. However, this possibility has desirable incentive properties because it can create an incentive to exert effort for self-interested bank managers who are, put simply, interested mostly in the continuation of their jobs. This is reasonable when the private benefits of control loom large, as may well be the case in emerging market economies with a weak institutional structure. In general, short-term debt has a disciplining effect in the presence of moral hazard. Indeed, in the extreme, the repayment of long-term debt may not be enforceable, and payment to the creditor may be ensured only by the threat of liquidation in an interim period (Bolton and Scharfstein 1990; Hart 1995).<sup>4</sup>

How does the theory relate to the trend in the banking industry of developed economies moving from the traditional business of taking deposits and granting

loans to the provision of services to investors (investment funds, advice, and insurance) and firms (consulting, insurance, mergers and acquisitions, underwriting of equity and debt issues, and risk management)? Banking in industrial countries is in a process of transformation (more advanced in the United States than in Europe), where the financial margin makes way for fee and commission revenue. Indeed, the share of assets held by banks relative to nonbank intermediaries is declining in developed economies (Allen and Santomero 2001). In contrast, in emerging market economies the traditional role of banks remains central.

In summary, in emerging market economies the traditional function of banks is all the more important, because financial markets are less developed and asymmetric information problems are more acute.

### *Fragility, Regulation, and the Safety Net*

The inherent fragility of the banking system, with asymmetric information at its root, leads to the failure of institutions, panic, and systemic crises that can have a major impact on the economy. The great depression of the 1930s is a good example, and more recent episodes of financial crises in the United States, Scandinavia, Mexico, East Asia, and Russia remind us of the potential for economic disruption. The failure of a bank has adverse consequences on nonfinancial firms precisely because individual bank–firm relationships are valuable (Petersen and Rajan 1994). In fact, even a contraction of bank capital may result in a credit crunch, with severe disruption to the private sector. This is especially evident in an emerging market economy, which is more dependent on the intermediary services of banks.

At the base of the fragility of banking is the coordination problem of depositors, who may decide to call back their short-term deposits and make a sound bank fail. The literature has two views of crises: the multiple-equilibrium panic view (Diamond and Dybvig 1983) and the information-based view (Gorton 1985, 1988; Jacklin and Bhattacharya 1988). According to the first, runs are triggered by events unrelated to the fundamentals, whereas according to the second, runs are triggered by bad news about the assets of the bank. Recently, those views have been reconciled by introducing asymmetric information and linking the probability of a run to the strength of fundamentals (Morris and Shin 2000; Rochet and Vives 2004; Goldstein and Pauzner 2005).<sup>5</sup>

Thus, a solvent bank may be subjected to a purely speculative panic, with depositors withdrawing funds and the bank being forced to quickly liquidate assets at a high cost. The cause of the problem is the dependence of banks on short-term debt (or the standard deposit contract). In addition, there is the danger of systemic risk owing to contagion from the failure of one entity, which may give rise to a strong negative externality both for the financial sector and for the real sector of the economy. For example, the failure of one entity may, through interbank market commitments,

lead to the failure of others (Allen and Gale 2001). Similarly, large variations in the price of assets, such as an abrupt fall in the stock market or the failure of a main intermediary, may generate a domino effect and systemic crises affecting the payment system.

In general, competitive banking will be excessively fragile, and the lender of last resort facilities and prudential regulation (discussed in the following section) will come to the rescue. The aim of regulation has been to provide the banking and financial systems with stability to avoid the negative effects associated with failing institutions and systemic crises. Other aims have been to protect the small investor and to promote the competitiveness of the system.

### *The Lender of Last Resort and the Policy Commitment Problem*

In industrial countries the lender of last resort and deposit insurance are basic to the stability of the banking system. There is a tendency, however, to protect banks and depositors above the levels required by the deposit insurance, in particular, under the too-big-to-fail policy. One reason is the potentially systemic consequences of the failure of a large institution, but more often help is a reflection of a time-inconsistency problem in the presence of a moral hazard problem. A well-intentioned lender of last resort—the central bank—will find it optimal ex post to help whenever this salvages the value of projects, whereas bankers, anticipating the help, will tend to exert sub-optimal effort, creating a moral hazard situation because the central bank is unable to observe the banker's level of effort in monitoring projects.

The time-inconsistency problem faced by a central bank arises because ex ante the central bank may want to commit to closing the bank if the returns are low (signaling a solvency problem), whereas helping the bank if the returns are only moderately low (signaling a liquidity problem). Such a commitment provides incentives for bank managers to monitor the projects they finance. In this way the central bank may implement the second-best solution in a competitive banking system.<sup>6</sup> However, ex post, costly liquidation of the projects may not be optimal, so the central bank may hesitate to carry out its threat. The commitment problem is compounded by the interest of a bank manager in the continuation of the bank. Building a central bank with a "tough" reputation can alleviate the time-inconsistency problem.

This commitment problem because of intertemporal inconsistency is aggravated in emerging market economies where institutions are weak and suffer from a lack of credibility and independence. It is difficult for central banks to build a reputation for disciplining banks because the central banker's effective horizon is short because of political instability. For example, in Argentina in the 1980s the average term in office for a central bank governor was less than a year although the legal term was four years (see Cukierman 1992, chapter 19). A related problem is the lack of legal protection for bank supervisors who attempt to impose discipline (again as in Argentina).

Then, even if the perceived problem is serious, the bank may be allowed to continue or even be granted help (World Bank 1998). A weak institutional structure allowing the regulated to unduly influence the regulators (regulatory capture) also explains why failure does not lead to a change of management. "Crony capitalism," where the government helps firms that are considered friends, is an extreme form of capture.

The consequence of the intertemporal inconsistency and regulatory capture is that the central bank of an emerging market economy that has a commitment problem will have incentives to induce inflation to reduce the real value of nominal debt commitments when banks or entrepreneurs are in trouble. This will avoid project liquidation, but it destroys incentives to exert effort and, in turn, devalues foreign investments in domestic currency. The outcome is a lack of foreign investment. As argued below, foreign-denominated short-term debt may be crucial for the access of an emerging market economy to international capital markets because it protects foreign investors from the devaluation of their claims by the actions of the government.

### *Prudential Supervision*

The too-big-to-fail policy, deposit insurance, and, in the extreme, blanket protection and bailouts introduce distortions into the decisions of financial entities. They reduce the incentive of depositors to monitor bank performance and, coupled with the bank's limited liability, they give rise to excessive risk taking. Bailouts eliminate the disciplining effect of closures and exacerbate risk taking and inadequate monitoring by bank managers.

The need for regulation is particularly acute when charter values<sup>7</sup> are low (and therefore incentives to take risk are high), and the social cost of failure is high (and therefore banking failure has a large impact). With either high disclosure requirements or risk-based insurance, banks pay for taking more risk, and capital requirements have a chance to be a sufficient instrument for controlling risk taking (table 1; see Diamond and Rajan 2000 for the role of bank capital in controlling excess fragility). Otherwise, capital requirements may need to be complemented with restrictions on the investments of banks to check risk taking.

Capital requirements together with supervision and market discipline are the three pillars of bank regulatory reform. The general trend in bank regulation is to check risk taking with capital requirements and appropriate supervision. Both risk-based (deposit) insurance and disclosure requirements have been proposed to limit risk-taking behavior (top and the bottom rows of table 1). Developed economies have tended to move in that direction. This movement has been accompanied by a reform of the 1988 Basel Accord on capital requirements to better adjust them for risk (Basel II). Basel II contemplates that banks can adopt either a "standardized" approach to capital requirements in which external rating agencies set the risk weight for different types of loans (say corporate, banks, and sovereign claims) or an

**Table 1.** Possible Banking Regimes, the Incentives to Take Risk, and the Necessary Regulatory Instruments When Charter Values are Low and the Social Cost of Failure Is High

<i>Banking regimes</i>	<i>Risk-taking incentives</i>		<i>Regulation</i>
	<i>Liability (rates)</i>	<i>Asset (investment)</i>	
Free banking, observable risk/high disclosure	Medium low	Absent	Capital requirements
Free banking, unobservable risk/low disclosure	Medium high	Maximal	Capital requirements and asset restrictions
Risk-insensitive insurance	High	Maximal	Capital requirements and asset restrictions
Risk-based insurance	Low	Absent	Capital requirements

*Source:* Based on Matutes and Vives (2000), Cordella and Yeyati (2002), and Hellmann, Murdock, and Stiglitz (2000).

internal rating–based approach in which banks estimate the probability of default (and also the loss, given default, in an advanced version of the method). The idea is to calibrate the capital requirement so that it covers the value at risk (expected and unexpected) from the loan under some assumptions.

It must be noted, however, that transparency has its limitations. Whereas it is feasible to introduce the disclosure requirements of the market positions of banks, it is more difficult to assess the risk level of the illiquid loan portfolio of a bank. Furthermore, more disclosure may induce information-based runs of investors, generating instability.

### *Regulation in an Emerging Market Economy*

If it is feasible to introduce risk-based insurance and disclosure requirements that eliminate moral hazard, capital requirements (risk-based themselves) may be a sufficient instrument to check risk taking and improve welfare (table 1). However, the characteristics of an emerging market economy cast doubt on the feasibility of such a strategy.

First, an emerging market economy is likely to face considerable uncertainty in terms of high economic volatility, high direct and indirect exposure to exchange rate risk, high maturity and currency mismatch, and high nondiversifiable risk in a typical loan portfolio. Two sources of increased risk are the high proportion of debt in foreign currency and of debt of short maturity (the reasons are explained later). The risk of high (and variable) inflation is, in a first instance, at the base of the use of those instruments. Furthermore, higher levels of risk are hidden behind the (false) security of a pegged exchange rate. For example, a bank feels protected because it has matched a dollar liability with a dollar-denominated loan without realizing that if the borrower earns income in pesos, a collapse of the peso will provoke a default. The currency

match has hidden credit risk. If the bank does not match a dollar deposit with a dollar-denominated credit, it becomes exposed directly to exchange rate risk.

Second, as noted, financial markets are less developed, and the monitoring role of intermediaries is enhanced in emerging market economies. The production of information on private sector activities and the general contracting environment are problematic. The background of these problems is the lack of economies of scale in the production of information and severe moral hazard and adverse selection problems. This implies, at the same time, that markets are thin and that the generation of information and contract enforcement roles rely relatively more on financial intermediaries. A further consequence of the thinness of financial markets is that banks have a high exposure to public debt (for example, government bonds) and are therefore vulnerable to inflationary strategies of the government and may be less able to match long-term investments by issuing appropriate liabilities (maturity mismatch). Short-term debt leaves banks and firms exposed to sharp increases in interest rates in response to a currency devaluation. At the same time a bank may have a harder time diversifying its portfolio because default probabilities may have a high correlation across projects. An obvious case is the collapse of the exchange rate.

All of these factors point to a riskier environment for banks. A currency crisis leads to a financial crisis and to strong effects in the real sector. A depreciation of the currency leads to a deterioration of balance sheets for firms and banks and to a decline in the net worth of the private sector. Because of asymmetric information problems, this decline in net worth will lead to a credit crunch, and banks with weak balance sheets will cutback on lending, exacerbating moral hazard and adverse selection problems (see Bernanke and Gertler 1989 for the general mechanism and Mishkin 1999a, 1999b for an application to the Tequila crisis in Mexico). A weak banking sector can also lead to a currency crisis. Kaminsky and Reinhart (1999) found that banking sector trouble typically precedes a currency crisis and that the currency crisis aggravates the banking crisis in a self-reinforcing manner. Finally, underdeveloped financial markets and no sound contracting environment increase the social cost of failure and the liquidation of projects. This means that the real effects of financial crises are multiplied.

Third, an emerging market economy will tend to have a weak supervisory structure. The reasons are rooted in the same factors that keep financial markets underdeveloped: the difficulty producing information and enforcement problems aggravated by the lack of protection for supervisors. Supervisors are either more easily corrupted, because of the lack of resources and low salaries, or more vulnerable to retribution if they do not acquiesce to the demands of lobbies, because of the lack of effective legal protection. Symmetrically, some banks may see how expropriatory regulatory decisions are made, perhaps because of their weaker political position, and this will induce a high rate of discounting. The consequence will be low charter

values as part of the bank profit is diluted, and this will happen even with a relatively low level of competition. The consequence will be enhanced incentives to take risk.

The characteristics of an emerging market economy of high uncertainty, increased likelihood and incidence of financial and currency crises, predominant financial role of banks, and weak supervisory structure lead to the policy conclusion that the regulatory strategy needs to be adapted to these conditions and to protect the fundamental role that banks play. Indeed, these characteristics make it much more difficult to follow the industrial country regulatory strategy in an emerging market economy. First, because information problems are more acute and the production of information is more problematic, it is more difficult to move toward a disclosure strategy. Second, risk-based deposit insurance can work only when insurance can be priced according to objective indicators of bank risk. Those indicators may be more difficult to obtain in an emerging market economy (they are difficult to obtain even in a developed economy). This makes the move toward a risk-based insurance strategy more difficult. Furthermore, the application of Basel II criteria to emerging market economies may be problematic. As Powell (2001) noted, these economies will have difficulty implementing the internal rating approach, especially because the new standards have not been calibrated for the environment in these countries, and so they will tend to adopt the standardized approach. But with the limited number of rated institutions in emerging market economies, this will mean little change from the current situation. Yet the problem of building a better link between risk and capital is, if anything, more acute there.

The corollary is that the regulation of banking and financial markets must be adapted for emerging market economies. Reliance on transparency and disclosure requirements as well as risk-based insurance and capital requirements is limited. Capital requirements in particular will need to be adapted to the conditions of emerging market economies (for example, public debt is risky because of inflation) and, most likely, complemented by other restrictions on the activity of financial institutions. At the same time, because of the high cost of the liquidation of projects and the social cost of failure in emerging market economies, competitive pressures and market discipline should not be set at the same level as in developed economies.

A broader consequence of the weak institutional structure of emerging market economies is that the policy commitment problem becomes central for attracting foreign capital. This problem and how external discipline can help overcome it are discussed in the following section.

Argentina coped with a policy commitment problem by importing external discipline through its currency board's "hard peg" (a fixed exchange rate backed by foreign reserves) and adopted a market discipline model (high disclosure levels, subordinated debt, limited deposit insurance, and risk-based capital requirements). Argentina was following a modern industrial country strategy anchored in the currency board (it was moving toward the top line of table 1, free banking with disclosure). The obvious

question is whether this model is feasible once the currency board has collapsed and the anchor has disappeared. The banking and currency crises were derived from the nonsustainability of the currency board in a recessionary context and with the underlying problems in Argentina, such as the lack of the credibility of institutions and the protection of property rights. The crisis was more like an informed run (such as that after the Tequila crisis) than a crisis derived from the moral hazard of bank managers or a coordination failure of the expectations of depositors. The financial system was reasonably well regulated and the currency board, when credible, provided discipline, thereby limiting excessive bailouts (see Calvo, Izquierdo, and Talvi 2003 for further details of the crisis).

## The Policy Commitment Problem and External Discipline

The government of an emerging market economy may devalue the claims of foreigners in domestic currency to protect the domestic private sector. This lack of policy commitment capacity derived from a weak institutional structure is a central problem for an emerging market economy that needs access to the international capital market. In a similar vein Tirole (2002) argued that the central market failure in external borrowing for an emerging market economy is the lack of contracting capacity between its government and foreign investors.

It has already been shown how short-term debt has a disciplining effect when there is a moral hazard problem, whether it is from the side of the entrepreneur seeking credit to finance a project or the banker monitoring a loan. This gives scope for a central bank in a competitive banking system to provide help in a range of returns while maintaining incentives by denying help when returns fall below the optimal critical threshold. The problem arises, however, that while ex ante it is optimal to commit not to help when returns are low, ex post, once effort decisions have been made, it is optimal to help avoid costly liquidation. A central bank with no policy commitment capacity will have incentives to induce inflation to reduce the real value of nominal debt commitments when banks or entrepreneurs are in trouble. This will avoid liquidation but destroy incentives to exert effort and discourage foreign lenders, who may see their claims devalued.

But to implement the second-best solution, a central bank must be able to commit not to help when returns are low. External discipline may come to the rescue (Vives 2002). An extreme form of importing external discipline, examined here for illustrative purposes, is "dollarization."

### *Dollarization*

Dollarization represents a commitment to a limited use of the lender of last resort facilities. Dollarization means that banking contracts are in "real" (dollar) terms. In a

dollarized regime help for the banking system (bailouts) must be arranged in advance, through stabilization funds or tax schemes, or precontracted in the international market. For example, when Argentina adopted a currency board in 1991–2001, it also established a contingent liquidity facility with international banks. Banks were also required to meet new liquidity requirements and to hold excess reserves, because the Convertibility Law of 1991 and the Charter of 1992 severely restricted the central banks' lender of last resort activity (Calomiris and Powell 2000). Most currency boards have established limited lender of last resort facilities. A stabilization fund can provide liquidity when needed, but it can be diverted (as happened in Mexico and Thailand).

Dollarization represents a commitment because it is costly to reverse. A currency board or a hard peg, as Argentina's recent experience suggests, represents a lesser commitment. Typically, the currency board is established by law (as in Argentina) and therefore can be dissolved by another law, which does, however, raise the cost of getting rid of the arrangement.

What are the costs and benefits of dollarization in a small open economy? In a competitive banking environment, projects are liquidated when the returns of the bank cannot cover the promised payment to depositors. However, in this competitive banking solution, there is typically excessive liquidation. The liquidation threshold imposed by a competitive banking environment is stricter than the second-best threshold because of risk sharing (a competitive bank has to promise a relatively high return to depositors that want to withdraw early to insure themselves although returns are higher in the long term) or because of a coordination failure among investors that induces a run on the bank. The excessive liquidation problem may be particularly acute if the coordination problem of investors is severe (Gale and Vives 2002).

The result is that dollarization provides a commitment not to help at the cost of not helping in circumstances where it would be optimal to do so *ex ante*. Therefore, dollarization is a third-best solution. Refusing help to the private sector creates incentives for managers to exert effort, but the probability of default implicitly chosen by the market may be much higher than the incentive-efficient level required to encourage managerial effort.

Partial dollarization can be a risky strategy. A bank may acquire dollar-denominated debt, accepting domestic dollar deposits or issuing dollar-denominated certificates of deposit in the international market. This opens the door to partial dollarization. In theory such an arrangement could deliver an optimal allocation of risk while providing incentives for bank managers to exert effort (Gale and Vives 2002). In practice, however, partial dollarization suffers from credibility and instability problems. The convertibility of dollar deposits by domestic residents can be suspended in a crisis. Furthermore, once banks accept dollar deposits, they are likely to make dollar-denominated loans to domestic residents (to avoid exchange rate risk and because they have a comparative advantage serving domestic customers). Credit risk will then resurface as residents earning the domestic currency have trouble paying the

dollar-denominated loans if the domestic currency depreciates. In fact, in trying to help by inducing inflation, the central bank may cause a devaluation of the currency, bankrupting the dollar-indebted households. The problem may be compounded by self-fulfilling expectations of a banking crisis.

### *Foreign Short-Term Debt*

Another way of importing external discipline is by acquiring foreign short-term debt. Short-term debt denominated in domestic currency will lose its disciplining effect if the central bank can devalue the currency through inflation when trouble arises. This is not the case with short-term debt denominated in foreign currency. Entrepreneurs in the small open economy may want to commit to high effort by borrowing in foreign currency. They will do so if the lower interest rate obtained with foreign debt more than compensates for the expected cost of liquidation (Jeanne 2000). Similarly, short-term debt ameliorates the default problem of an opportunistic populist government that repays debt only if it is in its short-term interest to do so and if default is costly. Taking foreign short-term debt may be good in a country prone to populist governments (Rodrik and Velasco 1999).

Consider an emerging market economy that has acquired foreign-denominated short-term debt to alleviate the moral hazard and commitment problem to access the international capital market. This short-term debt serves to discipline the private sector to exert effort to improve returns and to prevent the government from devaluing the claims of foreign investors. This allows access to the international capital market in amounts beyond the country's reserves (Although the borrowing capacity of the emerging market economy country is still limited by its international collateral [Caballero and Krishnamurthy 2002,2003]).

When returns are below a certain threshold, projects should be restructured to provide incentives to exert effort. The market may result in excessive liquidation of entrepreneurial projects, because of the coordination failure of investors, for example. This will always happen when the moral hazard problem in the country is moderate (implying a low restructuring threshold for returns). In this case, and to preserve incentives, projects should not be foreclosed when the country is insolvent but not too close to the solvency threshold. Instead, the market will foreclose except when the country is well above the solvency point. However, when the moral hazard problem is severe (implying a high restructuring threshold for returns), a solvent country may need to restructure when returns are close to and above the solvency threshold. Then the coordination failure of international investors may be a blessing in disguise because it may enforce the optimal restructuring threshold. That is, it may (although it need not) provide optimal incentives ex ante (Rochet and Vives 2004).

Foreign short-term debt is a two-edged sword. It works as a disciplining device but leaves the door open to systemic risk, speculative attack, and a sudden reversal of

capital flows. This means that in a country that faces a moral hazard problem over-exposure to foreign short-term debt is possible but that in general its optimal amount will not be zero provided that there is no extreme moral hazard problem. Obviously, the optimal amount of foreign-denominated debt will also depend on the size of the tradable sector in the economy, which determines the amount of effective international collateral that the country can pledge.

Another reason to issue foreign short-term debt is risk sharing when foreign investors are risk-averse. This is akin to the risk-sharing role of the standard deposit contract that a bank offers to its risk-averse depositors. In this case the issuing countries absorb part of the risk and can obtain more favorable terms with respect to expected repayment than borrowing long term. Broner, Lorenzoni, and Schmukler (2004) provided a model where the maturity structure of the debt arises as the resolution of a risk-sharing problem between the government and bondholders. The government of an emerging market economy lowers the probability of a liquidity crisis, shifting risk to bondholders, by issuing long-term debt. This risk is reflected in equilibrium in a higher risk premium and cost of borrowing. In consequence, the government faces a tradeoff between safer long-term borrowing and cheaper short-term debt. Broner, Lorenzoni, and Schmukler also provided evidence of this effect, particularly in times of crisis when the term premium increases and issuance shifts toward shorter maturities.

Empirical results are mixed. Sachs, Tornell, and Velasco (1996) and Francis and Rose (1996) found either weak evidence or no statistical effect of short-term debt on crises. Eichengreen and Rose (1998) found that higher levels of short-term debt decrease the probability of banking crises. Radelet and Sachs (1998) and Rodrik and Velasco (1999) found that the ratio of short-term debt to reserves is a robust predictor of financial crisis (sharp reversal of capital flows).<sup>8</sup> Finally, Detragiache and Spilimbergo (2004) found that short-term debt increases the probability of a debt crisis (although, obviously, this could also be interpreted as more fragile countries having to borrow short term).

A potential alternative to external discipline is financial indexation to some domestic price level. This has been tried in Chile with the Unidad de Fomento (unit of account). The problem is that any index is bound to be imperfect when the economy experiences large swings in relative prices (Calvo and Talvi 2005).

In summary, external discipline through a relatively large amount of foreign short-term debt will be good in emerging market economies or developing countries with a policy commitment problem in which (a) there is a significant but not extreme moral hazard problem, (b) monitoring effort by bankers and entrepreneurs is important for improving returns, and (c) the cost of liquidating projects is not large.

The optimal amount of foreign short-term debt will presumably increase with the size of the tradable sector and the effectiveness of effort by managers in improving returns and decrease with the cost of liquidating projects.

Candidates for external discipline are countries with a long way to go in political stability, rule of law, contract enforcement, and institutional development and supervision, and those that rely on bank monitoring to make finance available for entrepreneurial projects. At the same time the moral hazard problem cannot be hopeless, and liquidation costs cannot be high. The analysis points to an intermediate range of countries with a weak institutional structure but without an extreme agency problem. Politically stable countries with a modern institutional structure and deep financial markets need not import external discipline.

## Assessing the Tradeoffs of External Discipline

For a range of emerging market economies, this section summarizes some rough indicators, based on Gale and Vives (2002), of the policy commitment problem and the determinants of the welfare analysis of importing external discipline: moral hazard, importance of effort for returns, and cost of liquidating projects.

The economies considered are Argentina, Brazil, Chile, Colombia, Ecuador, Mexico, Peru, Uruguay, and Venezuela in Latin America; Hong Kong (China), Indonesia, Malaysia, Philippines, Singapore, the Republic of Korea, Taiwan, and Thailand in Southeast Asia; and Turkey. Attention focuses on Latin America and Southeast Asia (although Turkey, on the periphery of Europe, is also considered).

An indicator of the extent of the commitment problem is given by the credibility of the central bank. A combined index (see Cukierman 1992, table 21.1) of legal independence, actual turnover of the central bank governors, and a response by experts to a questionnaire can be used to assess the overall credibility and reputation of a central bank. The combined index (weighting the three indexes by their relative contribution in explaining the variations in the rate of depreciation of the value of the currency) provides a ranking of central bank independence of 46 countries during the 1980s. At the bottom are economies such as Argentina, Brazil, Chile, Turkey, Venezuela, Mexico, Peru, Uruguay, the Republic of Korea, and Indonesia. Above these are the Philippines and Thailand. However, the Philippines has the same level of overall central bank independence as Kenya, and Thailand has the same level as Greece. Developments since the 1990s may qualify the situation for some countries (for example, Chile). If we only look at the actual turnover of governors (for the period 1950–89) as an index of credibility, the situation is not different (see Cukierman 1992, table 19.4).

The severity of the moral hazard problem for the bankers or entrepreneurs can be proxied by a battery of legal indicators reflecting the rule of law in different countries (LaPorta and others 1998). These include the indicators of the efficiency of the judiciary system, the law and order tradition of the country, corruption, the risk of expropriation, and the risk of contract repudiation by the government, as well as a

rating on accounting standards. Low marks on these indicators would suggest a severe moral hazard problem for the banker. For example, in countries with poorly defined or poorly defended property rights, some of the bank's marginal returns will be captured by others, which increases the marginal cost of ensuring high returns. Gale and Vives (2002, table 1) classified the marks given in La Porta and others (1998) as low, medium, and high. Under this classification, Peru, Indonesia, and the Philippines (and perhaps Venezuela and even Uruguay) have a severe moral hazard problem. Argentina, Brazil, Chile, Colombia, Ecuador, Mexico, the Republic of Korea, Taiwan (China), Thailand, and Turkey have a significant but not extreme moral hazard problem. Hong Kong (China), Malaysia, and Singapore have a moderate moral hazard problem.

There is also a link between the (indirect) moral hazard indicators (related to the rule of law and accounting standards) and the importance of effort by the banking manager to obtain returns. In countries with a severe or significant moral hazard problem with the firms in the private sector, suggested by low marks in the rule of law indexes, effort by the bank manager to obtain returns will also be important. This means that moral hazard is bound to be a problem in the set of economies considered here, perhaps with the exception of Hong Kong (China), Malaysia, and Singapore. The importance of monitoring efforts by the private sector (bank managers or entrepreneurs) to obtain project returns could also be proxied by the importance of banks in the financial system. This can be measured by the amount of bank assets to total financial assets. In the economies of the sample, only the Republic of Korea is below a ratio of 50 percent (data are unavailable for Hong Kong, China, Indonesia, and Singapore).

The cost of liquidation of projects can be approximated by the level of rights of creditors (high creditor rights are linked to a lower cost of liquidation of projects) and the level of development of the financial system (more developed financial systems can cope better with adverse selection problems at the root of costly liquidation). The level of development of the financial system can be approximated by the ratio of credit to the private sector to GDP, the ratio of stock market capitalization to GDP, and accounting standards. A combination of those indicators yields the following classification for the cost of liquidation:

- High: Peru, Uruguay, and Venezuela.
- Medium high: Argentina, Brazil, Colombia, Mexico, and Turkey.
- Medium low: Ecuador, the Philippines, and Indonesia.
- Low: Chile, Hong Kong (China), Singapore, the Republic of Korea, Malaysia, Thailand, and Taiwan (China).

It is worth noting the regional difference between Latin America (with the exception of Chile) and East Asia. This would be consistent with the better-than-expected recovery of Southeast Asia from the last financial crisis.

The indicators point to a middle range of economies where external discipline can be a good idea. Those economies are at the intersection of a significant or moderate moral hazard problem with medium-low or low cost of liquidation. Among them are some Latin American countries, such as Chile and Ecuador, as well as some East Asian tigers, such as Hong Kong (China), Malaysia, the Republic of Korea, Singapore, Taiwan (China), and Thailand. Peru, Indonesia, and the Philippines (Venezuela and even Uruguay) are not included because of the severity of the moral hazard problem, to which a potentially high cost of project liquidation is added for Peru, Uruguay, and Venezuela. A medium-high cost of project liquidation is found for Argentina, Brazil, Colombia, and Mexico, as well as Turkey. Those would be candidates for a high dose of external discipline because of the potential benefit to be derived, but the cost would also be high. At the same time Hong Kong (China), Malaysia, and Singapore would derive less benefit from external discipline because the monitoring of projects is relatively less important and the moral hazard problem less severe.

## A Catalog of Proposals

Various proposals are frequently made to solve the banking and regulatory problems of emerging market economies. This section briefly surveys some of these—narrow banking, monetary union, foreign banks, public banks, and offshore banking—and argues that none of them is a panacea that can provide the desired solution.

### *Narrow Banking*

Narrow banking has been proposed to solve the instability problem and to limit the exposure of insurance funds and taxpayers to banking trouble. The extreme of a narrow bank is one that holds 100 percent reserves. Friedman (1959) advocated such banks (and proposed that reserves be remunerated). Other proposals would have the bank invest only in safe short-term securities (such as U.S. Treasury bills). Deposit insurance would then serve only to protect against fraud, and nonbanks would not be insured. In this scheme a financial holding company could have two subsidiaries: a narrow bank arm (mutual fund) and a finance company arm that would grant loans and fund them with uninsured liabilities (Litan 1987). Other proposals would allow the bank to take more risk.

The basic idea is to convert the bank into a mutual fund that provides liquidity and dominates the traditional bank in the deposit dimension. Indeed, in a mutual fund where the investor is promised a share of the value of the fund, there is no incentive to run. Deposit insurance is unnecessary (except perhaps to protect against fraud) and need not generate moral hazard. The case for the narrow bank is buttressed by the increasing process of securitization of credits, by some studies that

conclude that economies of scope between deposit and lending are small (Pulley and Humphrey 1993), and by advantages in the control of the money supply (however, the last two conclusions are disputed by Benston in Fernandez and Schumacher 1997).

With sufficiently developed financial markets, the criticisms that a narrow bank may not have enough safe assets available to invest or that there would be insufficient funds left to finance long-term projects are unlikely to hold, but they do have force in emerging market economies with an underdeveloped capital market. However, there is a deeper criticism of the proposal: narrow banking throws away the baby with the bath water. That is, it destroys the financing of illiquid loans with liquid liabilities (see Wallace 1996 for a version of this argument). These are mostly business loans, which are opaque and nonsecuritizable, in contrast to mortgages and consumer credit. As already shown, the fragile structure of banks allows them to create liquidity for business. Indeed, there is evidence that deposit taking and loan provision go together. Nakamura (1988) found informational economies of scope between the two functions. For example, often the first thing a bank requires before giving credit is that the firm open an account so that the bank can monitor the firm's operations. In fact, one can regard deposit taking and credit lines as essentially the same operation: liquidity provision on demand. Imperfect correlation in these activities leads then to synergies (empirically documented by Kashyap, Rajan, and Stein 2002).

Most narrow banking proposals just push the fragility problem plus liquidity creation for firms to another level. For example, finance companies (for long-term projects) will provide liquidity financed by negotiable certificates of deposit, but this retains the coordination problem of investors. This means that narrow banking does not address the basic problem for which safety nets are intended: the adverse consequences of banking fragility in terms of credit supply and externalities for the private sector of the economy. A potential implication is that the commitment not to insure finance companies that take on the functions of banks may not be credible. In fact, one may believe the opposite. Narrow banking may encourage discretionary bailouts because it limits explicit insurance to the narrow bank, which is politically unrealistic given the need of a safety net for the banking system.

In conclusion, the narrow banking proposal, in its strict form, does away with the main function of banks and, therefore, will probably be particularly detrimental to emerging market economies, where financial markets are less developed and securitization is less advanced. In contrast, narrow banking in its softer forms may be appropriate for economies with well-developed financial systems such as in the United States, where liquidity is abundant. In fact, money market funds are providing more and more transaction services in the United States. Put another way, emerging market economies have a greater need for bank monitors to finance and elicit returns from entrepreneurial projects. It is also true that emerging market

economies with a market discipline problem may need to place some asset restrictions on banks' balance sheets to control risk taking. Restrictions may be minimum diversification levels, limitations on concentration of risks, growth capacity of loans in new areas, and so on. This goes some way toward the narrow banking proposal, in the sense of restricting risk taking on the asset side, but without terminating the financing of illiquid loans with deposits.

At the same time there are more limited forms of narrow banking that allow informational economies of scope between deposits and loans to be preserved. For example, banks could be allowed to offer narrow bank accounts, essentially money market funds insured only against fraud, on top of their "standard" uninsured deposits and loan provision facilities (Mishkin 1999a, Mishkin 1999b). This proposal apparently limits the cost of deposit insurance, but incentives would still remain to bail out the non-narrow part of the bank because of the adverse consequences of a failure for the economy.

It may help think counterfactually what would have happened if narrow banks had been in place in Argentina during the currency board regime. Would the banking crisis have been averted? Probably not. Consider first the orthodox form of a narrow bank as a separate institution investing deposits in safe short-term assets (such as U.S. Treasury bills). In this situation there would be a run only if there is a fear of confiscation (for example, with forced conversion of dollar deposits into depreciated pesos). The rational reason for the fear is that with an unsustainable currency board, banks and finance companies that provide loans to firms and consumers will be in trouble anyway. Narrow banking may make this confiscation more difficult, or more evident, but may not be able to avoid it. In a crisis situation with weak property rights, the pressure to capture the sound assets of the narrow bank may just be too great. A softer form of narrow banking with narrow accounts in a multipurpose bank may be more vulnerable. However, even if depositors would not have run on narrow banks, the collapse of the currency board would have induced default on dollar-indebted agents and therefore brought down the financial institutions that lent to them. At best, the crisis is transferred to another level, which means that the cost of the crisis is shifted to the investors or depositors in the financial companies that lend to the private sector.

This discussion points to the fact that the first problem to be solved in an emerging market economy is, in many instances, political: to insure property rights so that depositors are confident that the contract with their bank will be respected and that the rules of the game are not changed in the middle of play. Narrow banking may help somewhat but need not solve this problem.

Second, an emerging market economy needs both an efficient payment and a transaction system and good facilities for credit that help sustain growth in the long term. The banking system should focus on both functions. This discards "strong" forms of narrow banking where deposit institutions are separated from

credit institutions. An emerging market economy needs the classical function of banks.<sup>9</sup>

### *A Monetary Union*

Another way to import discipline is by transferring some political sovereignty by forming a monetary union with other countries (such as Mercosur in trade, for example) and establishing an independent central bank. This solution is likely to be problematic, however, mainly because it is not clear how to build an independent central bank from the monetary union of countries whose central banks have shaky reputations. For example, the European Monetary Union (EMU) would probably not have arisen had not the Bundesbank had such strong anti-inflationary credentials.

Another issue is that the organization of the lenders of last resort in a monetary union with no central political authority is likely to be contentious as, again, the EMU shows. A decentralized organization for the lenders of last resort raises serious issues of effectiveness, whereas a centralized organization needs to clarify the fiscal issue (who will pay for institutions that end up insolvent) and determine in what situations the centralized lender of last resort will not help (a systemic problem in a small country may not be systemic in the larger area of the monetary union, for example). (See Vives 2001 for a discussion of those issues in the European context.) The experience of the EMU points to the difficulties that advanced economies with well-developed institutions face in establishing a common central bank. At the present stage in Latin America, the avenue of a monetary union does not seem realistic.

### *Foreign Banks*

It may be thought that with an external discipline and a banking system mostly in foreign hands, stability should be guaranteed because of the support provided by strong international banks supervised abroad and with the backing of their respective central banks. The Argentine crisis dispels this notion. While the headquarters of foreign banks would provide help to a subsidiary when a problem develops (because the brand name and the franchise value of the bank are at stake), this need not hold when a systemic problem develops. This is the more so if the systemic problem (due to the nonsustainability of the currency board arrangement) is compounded by confiscatory measures.

Furthermore, even if foreign bank headquarters were willing to help, they need not do so at the optimal social level, because they will not take into account the external effects of their help. For example, the headquarters of foreign banks may want to limit their exposure to a country that may face a currency crisis and therefore will tighten liquidity provision to branches or subsidiaries in the country. (There is an important distinction between a branch and a subsidiary: headquarters have to back the deposits in a branch but need not do it for a subsidiary.)

Furthermore, the incentives of the foreign lender of last resort and supervisor are not necessarily in line with local interests. A foreign supervisor will not take into account the consequences (systemic or not) for domestic residents of a restructuring of a local branch or subsidiary. The foreign supervisor will care only about the consequences for systemic stability at home of a crisis of a subsidiary abroad.

Foreign banks are thus no panacea for solving the lender of last resort and supervisory problems in an emerging market economy. Whereas successful in Panama, which has no lender of last resort facility except, seemingly, a large U.S. bank (Chapman 1999; Moreno-Villalaz 1999), Panama is not a typical emerging market economy.

Furthermore, some research using Argentinean data claims that large foreign-owned institutions concentrate on large projects and may have trouble providing relationship loans to opaque small firms (Berger, Klapper, and Udell 2001). If confirmed, this may call for an appropriate balance of foreign and local institutions or at least for the use by foreign banks of local subsidiaries rather than branches. A local subsidiary typically comes from the foreign bank's acquisition of a local bank and, therefore, at least part of the informational capital of the local bank should be incorporated into the new or transformed entity.

### *Public Banks*

There is an agreement that public banks have been the worst performers in Argentina according to standard criteria (see, for example, Berger, Klapper, and Udell 2001) and that they were the principal cause of the run in 2001. The question is what role should public banks have in an emerging market economy. A theoretical case could be made for their role in extending credit where the market fails because of acute asymmetric information problems. However, the case for privatization is strong once it is acknowledged that public banks are the main culprits behind the excessive bailout problem, with a destabilizing influence, and once the political economy issues are weighed in the balance. At the end of the day, the public sector cannot provide what the private sector has failed to provide.

Mandating a narrow bank charter for public banks limits their exposure but at the same time removes their only potentially redeeming value: providing loans to informationally opaque business. Private banks can provide subsidized credit to families and firms, as well as "universal service" for transaction accounts, if deemed necessary. The proper way to do this is to auction the service to well-capitalized private institutions. If the privatization route is taken, attention should be paid to ensure that prospective bidders have enough capability and local knowledge to provide the required relationship-based banking services so crucial to small business.

## *Offshore Banking*

Offshore banking consists of the provision of financial services and cross-border intermediation of funds to nonresidents by banks residing in offshore financial centers. Offshore financial centers are characterized by low levels of regulation and information disclosure and tend to be tax havens.

During the Tequila crisis offshore establishments of Argentine banks suffered a run parallel to the onshore run. Several offshore institutions and their onshore parents failed, and investors onshore were left in a weak position to claim assets, because the offshore institutions were in different jurisdictions (Errico and Musalem Borrero 1999).

As with narrow banking, with offshore banking the institutional credibility-confiscation problem is alleviated, in this case more substantially, but it is not completely solved. The question is whether offshore arrangements can be designed so that it becomes difficult to confiscate the monies once they reach national soil. A benefit of offshore banking is then to alleviate the confiscation problem. However, with offshore banking the benefit of a banking system close to business would be lost, and the additional cost of operation in a different jurisdiction should be added to this.

## The Role of an International Agency

There are two types of market failures that an international agency such as the IMF may help overcome. The first is the coordination problem faced by international investors in a crisis situation. The second is the commitment problem of domestic institutions such as a national central bank. The two are intertwined and, as we will see, the first may help alleviate the second. The reason is that an ex post coordination failure may help create ex ante incentives.

An international agency, combining the functions of crisis lender and crisis manager, may be helpful provided it is complemented with a policy of prompt corrective action and facilities for orderly failure resolution.

The IMF has typically relied on ex post policy conditionality for granting help to a country—that is, the need for a country to commit to follow certain policies to be helped after a crisis. Although the IMF can ask for collateral, it rarely does. The reason is that the IMF is seen as a preferred creditor: not repaying the IMF essentially means being denied access to the international capital market. At the same time the IMF has also played a role as crisis manager in negotiating the way out of crisis and helping to set up financial packages (Fischer 1999).

An international lender of last resort—whether helping countries in trouble much the way a central bank helps individual banking institutions, coordinating the

interventions of the three main central banks, or acting as a global central bank issuing an international currency — can follow a policy of injecting liquidity in international financial markets (see Eichengreen 1999 for a survey of the different proposals). The last approach is developed in several proposals that adapt Bagehot's doctrine to international lending. The *Report of the International Financial Institution Advisory Commission* (IFIAC 2000), chaired by Allan H. Meltzer, recommended that the IMF lend according to Bagehot's rules: lending at a penalty rate with collateral only to a set of countries that meet certain requirements in their banking systems (adequate capitalization of domestic banks, free entry of foreign banks) as well as having a commitment to fiscal standards and to reporting accurate financial information. A significant difference between the approaches is on the required size of the international lender of last resort (Jeanne and Wyplosz 2001). In the first case an issuer of international currency is needed, whereas in the second the intervention is bounded by the difference between the short-term foreign exchange liabilities of the banking sector and the foreign reserves of the country in question.

Consider the more realistic second approach. The main tension identified in the debate is between those who emphasize the crisis prevention effect of liquidity support (Fischer 1999) and those who are worried about generating moral hazard in the country being helped (IFIAC 2000). This reflects the tension between the need to provide ex ante incentives to the country to limit moral hazard and the need to provide help once the crisis occurs.

The conceptual framework is as described earlier for an emerging market economy that has acquired foreign-denominated short-term debt to alleviate the moral hazard and commitment problem to access the international capital market (Rochet and Vives 2004). When returns are below a certain threshold, projects should be restructured to provide incentives to exert effort. This second-best solution can be implemented with the help of an international agency. The following scenarios can be considered:

- *No bail-in, no bail-out.* With no international lender of last resort and interrupted access to international capital markets, country projects are liquidated whenever withdrawals by foreign fund managers are larger than foreign reserves. This limits investment.
- *Bail-in, but no bail-out.* With no international lender of last resort but access to international capital markets, some costly project liquidation is avoided with fire sales of assets, but excessive liquidation of entrepreneurial projects will still typically occur.
- *Bail-in and bail-out.* An international agency with crisis lender and crisis manager roles can complement country access to international capital markets to enforce this second-best solution. The appropriate policy should depend on the extent of the moral hazard problem in the country considered.

When the moral hazard problem in the country is severe, a policy of prompt corrective action whenever the country is close to insolvency is needed to complement the international lender of last resort facility. A solvent country may need to restructure when returns are close to the solvency threshold.

When the moral hazard problem in the country is moderate, in addition to the international lender of last resort help for a solvent country, an orderly resolution of the failure process or help is needed for an insolvent country that is close to the solvency threshold. This may be interpreted as a mechanism similar to the sovereign debt restructuring mechanism of the sort considered by the IMF with the objective of restructuring unsustainable debt.<sup>10</sup>

An international lender of last resort that follows a modified Bagehot's prescription complemented by prompt corrective action and orderly failure resolution facilities can implement a second-best solution provided that it is well informed about country returns. One possibility is that the international lender of last resort does in-depth country research and has supervisory knowledge of the banking system of the country where the crisis occurs. This international lender of last resort should also have statutory independence (along the lines of, for example, De Gregorio and others 1998) that allows building a reputation for not granting help when it is not *ex ante* efficient to do so and thereby not contributing to moral hazard. *Ex ante* conditionality in terms of the rule of law and accounting standards makes sense because they contribute to assessing the extent of the moral hazard problem in the country and, therefore, the appropriate policy.

Furthermore, the international lender of last resort should provide a loan only when the country has enough international collateral. The critical amount of collateral should be determined by the financial position associated with the return below which help should not be provided. Obviously, to determine the minimum required amount of collateral (and the range where help has to be provided), we need the supervisory knowledge of the economy. A minimum amount of collateral is established not to secure the loan but to impose a threshold below which help is not given.

## Policy Recommendations and Issues for Further Research

The underlying problems of regulation are the same in industrial countries and developing countries and emerging market economies. However, in developing countries and emerging market economies, the asymmetric information problems are more acute, and the institutions are weak. Several consequences follow from this.

First, regulation must be adapted because reliance on transparency and disclosure requirements as well as risk-based insurance is limited. This means in particular that capital requirements will need to be tuned to the emerging market economy context

and complemented by other types of restrictions on the activity of financial institutions. At the same time the cost of the liquidation of projects or the social cost of failure may be larger in developing countries and emerging market economies. This calls for moderating competitive pressures and market discipline.

A second consequence is that the policy commitment problem of domestic institutions becomes central. From this it follows that developing countries and emerging market economies need to import external discipline. This can be accomplished by acquiring foreign short-term debt. The analysis provides rough guidance on the optimal amount of external discipline that is good for a country as a function of the severity of the moral hazard problem and the cost of liquidating projects. There is a middle range of countries—where the moral hazard problem is significant but not extreme and the cost of liquidating projects is moderate—for which importing external discipline is worthwhile.

A corollary of the analysis is that developing countries and emerging market economies need differentiated treatment not only from developed economies but also among themselves, especially with respect to the levels of external discipline.

Finally, international institutions such as the IMF may be instrumental in helping emerging market economies overcome the commitment problem. However, several requirements need to be met: the institutions should be independent; they should lend conditionally on the institutional structure of the country (not on policy); and they should require collateral (not to insure repayment but as incentive).

Many issues need further research, including the optimal regulatory policy mix for every stage of development and the quality of the institutional structure; the interplay between the coordination problem causing fragility in the banking and financial system, and the moral hazard problem caused by bailouts in response to fragility; and the determinants of the optimal level of short-term foreign debt for a small open economy with a policy commitment problem.

## Notes

Xavier Vives is Professor of Economics and Finance at IESE Business School, University of Navarra, Barcelona, and Research Professor at Institutio Catalana de Recerca i Estudis Avançats-Universitat Pompeu Fabra (UPF); his email address is xvives@iese.edu. The author is grateful to the World Bank and the Inter-American Development Bank for support on projects related to this article as well as to project SEJ2005-08263 at UPF of the Directorate-General for Research of the Spanish Ministry of Education and Science. The article draws on the author's joint work in banking with Douglas Gale, Carmen Matutes, and Jean-Charles Rochet. The author is grateful to Ramon Caminal, Fernando Broner, Eduardo L. Yeyati, three referees, and Shanta Devarajan for helpful comments.

1. For the Argentina case Calomiris and Powell (2000, p. 1) argued that "the banking sector suffered from ineffective regulation and supervision and repeated, forced government rescues contributed significantly to Argentina's past fiscal and inflationary problems."

2. *Moral hazard* refers to the effects that insurance may have on the behavior of people with insurance, who may take greater risks than they would without it because they know they are protected, with adverse consequences for the insurance company. The problem arises because the insurance

company cannot observe the effort the agent makes to prevent a loss. In economic analysis moral hazard refers to hidden actions in a contractual relationship. *Adverse selection* arises when the insurer knows more than the company about the probability of the loss happening. It is an issue of hidden information. In general adverse selection involves the adverse consequences for uninformed parties of the actions of privately informed ones. These asymmetric information problems are documented for a range of emerging market economies later in the article.

3. The demand deposit contract, redeemable at par, either creates a coordination problem for investors that does not allow bankers to extort rents on their abilities to collect illiquid loans (Diamond and Rajan 2001) or disciplines bank managers subject to a moral hazard problem (Calomiris and Kahn 1991; Gale and Vives 2002).

4. Moral hazard may also arise from the side of the entrepreneur asking for credit to finance a project. If effort is devoted to the project, returns are improved. If the entrepreneur takes a loan that is forcibly short term, the threat of liquidation if the returns (which are not verifiable) are low and cannot cover the promised debt payment provides the incentive for effort in the project. Effort improves returns, but liquidation is costly.

5. Postlewaite and Vives (1987) provided an early model with a unique equilibrium, where the probability of a crisis is determined by the realization of the liquidity needs of depositors, which are private information.

6. In the presence of a moral hazard problem efficiency (of the second-best type because of incentive considerations) requires that the expected utility of investors and depositors be maximized subject to the constraint that the bank manager exerts effort. This can be accomplished by liquidating the project when observable interim returns are lower than a certain threshold (this is the minimal threshold that induces the manager to exert effort; a higher threshold would just increase the costs associated with liquidation). The issue is how to achieve or approximate this second-best outcome. This is discussed again in the following sections.

7. The charter value is the franchise value of the bank (forward-looking capitalized discounted expected profits). It can be measured by the market to book asset ratio. A high charter value means that the bank has a lot to lose if closed.

8. Rodrik and Velasco (1999) also found that a greater short-term exposure aggravates the crisis once capital flows reverse.

9. See Fernandez and Schumacher (1997) for a more positive view of narrow banking for Argentina.

10. Bolton (2003) discussed sovereign debt restructuring mechanisms from the perspective of corporate bankruptcy theory and practice.

## References

- Allen, F., and D. Gale. 2001. "Financial Contagion." *Journal of Political Economy* 108(1):1–33.
- Allen, F., and A. Santomero. 2001. "What Do Financial Intermediaries Do?" *Journal of Banking and Finance* 25(2):271–94.
- Berger, A. N., L. F. Klapper, and G. F. Udell. 2001. "The Ability of Banks to Lend to Informationally Opaque Small Businesses." *Journal of Banking and Finance* 25(12):2127–67.
- Bernanke, B., and M. Gertler. 1989. "Agency Costs, Net Worth, and Business Fluctuations." *American Economic Review* 79(1):14–31.
- Bolton, P. 2003. "Toward a Statutory Approach to Sovereign Debt Restructuring: Lessons from Corporate Bankruptcy Practice around the World." *IMF Staff Papers* 50 (Special Issue):41–71.
- Bolton, P., and D. Scharfstein. 1990. "A Theory of Predation Based on Agency Problems in Financing Contracting." *American Economic Review* 80(1):93–106.

- Broner, F., G. Lorenzoni, and S. Schmukler. 2004. "Why Do Emerging Economies Borrow Short Term?" Economics Working Papers 838. Department of Economics and Business, Universitat Pompeu Fabra, Barcelona.
- Caballero, R., and A. Krishnamurthy. 2002. "A 'Vertical' Analysis of Monetary Policy in Emerging Markets." MIT Department of Economics Working Paper 02-16. Cambridge, Mass.
- . 2003. "Excessive Dollar Debt: Financial Development and Underinsurance." *Journal of Finance* 58(2):867-93.
- Calomiris, C., and C. Kahn. 1991. "The Role of Demandable Debt in Structuring Optimal Banking Arrangements." *American Economic Review* 81(3):497-513.
- Calomiris, C. W., and A. Powell. 2000. "Can Emerging Market Bank Regulators Establish Credible Discipline? The Case of Argentina 1992-1999." NBER Working Paper 7715. National Bureau of Economic Research, Cambridge, Mass.
- Calvo, G., and E. Talvi. 2005. "Sudden Stop, Financial Factors and Economic Collapse in Latin America: Learning from Argentina and Chile." NBER Working Paper 11153. National Bureau of Economic Research, Cambridge, Mass.
- Calvo, G., A. Izquierdo, and E. Talvi. 2003. "Sudden Stops, the Real Exchange Rate and Fiscal Sustainability: Argentina's Lessons." NBER Working Paper 9828. National Bureau of Economic Research, Cambridge, Mass.
- Chapman, G. 1999. "La Experiencia de Dolarización en Panama." Paper Presented at the Conference on Opciones Cambiarias para la Región, Sponsored by the Inter American Development Bank, July 23-24. Panama City.
- Cordella, T., and E. L. Yeyati. 2002. "Financial Opening, Deposit Insurance, and Risk in a Model of Banking Competition." *European Economic Review* 46(3):471-85.
- Cukierman, A. 1992. *Central Bank Strategy, Credibility, and Independence: Theory and Evidence*. Cambridge, Mass.: MIT Press.
- De Gregorio, J., Eichengreen, B., Ito, T., and C. Wyplosz. 1998. *An Independent and Accountable IMF: Geneva Reports on the World Economy, 1*. Oxford: Information Press.
- Detragiache, E., and S. Spilimbergo. 2004. "Empirical Models of Short-term Debt and Crises: Do they Test the Creditor Run Hypothesis?" *European Economic Review* 48(2):379-89.
- Diamond, D., and P. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91(3):401-19.
- Diamond, D., and R. Rajan. 2000. "A Theory of Bank Capital." *The Journal of Finance* 55(6):2431-65.
- Diamond, D. W., and R. Rajan. 2001. "Liquidity Risk, Liquidity Creation and Financial Fragility: A Theory of Banking." *Journal of Political Economy* 109(2):287-327.
- Eichengreen, B. 1999. *Toward a New International Financial Architecture: A Practical Post-Asia Agenda*. Washington, D.C.: Institute for International Economics.
- Eichengreen, B., and A. Rose. 1998. "Staying Afloat When the Wind Shifts: External Factors and Emerging-Market Banking Crises." CEPR Discussion Paper 1828. Centre for Economic Policy Research, London.
- Errico, L., and A. Musalem Borrero. 1999. "Offshore Banking: An Analysis of Micro-and Macro-Prudential Issues." IMF Working Paper 99/05. International Monetary Fund, Washington, DC.
- Fernandez, R., and L. Schumacher. 1997. "Does Argentina Provide a Case for Narrow Banking?" In S. K. Bery and V. F. Garcia, eds., *Preventing Banking Sector Distress and Crises in Latin America: Proceedings of a Conference Held in Washington, D.C., April 15-16, 1996*. Discussion Paper 360. World Bank, Washington, DC.
- Fischer, S. 1999. "On the Need for an International Lender of Last Resort." *Journal of Economics Perspectives* 13(4):85-104.

- Frankel, J., and A. Rose. 1996. "Currency Crashes in Emerging Markets: An Empirical Treatment." *Journal of International Economics* 41(3-4):351-66.
- Friedman, M. 1959. *A Program for Monetary Stability*. New York: Fordham University Press.
- Gale, D., and X. Vives. 2002. "Dollarization, Bail Outs, and the Stability of the Banking System." *Quarterly Journal of Economics* 117(2):467-502.
- Goldstein, I., and A. Pawner. 2005. "Demand Deposit Contracts and the Probability of Bank Runs." *The Journal of Finance* 60(3):1293-327.
- Gorton, G. 1985. "Bank Suspension of Convertibility." *Journal of Monetary Economics* 15(2):177-93.
- . 1988. "Banking Panics and Business Cycles." *Oxford Economic Papers* 40(4):751-81.
- Hart, O. 1995. *Firms Contracts and Financial Structure*. Oxford: Clarendon Press.
- Hellmann, T. F., K. Murdock, and J. Stiglitz. 2000. "Liberalization, Moral Hazard in Banking, and Prudential Regulation: Are Capital Requirements Enough?" *American Economic Review* 90(1):147-65.
- Holmstrom, B., and J. Tirole. 1997. "Financial Intermediation, Loanable Funds and the Real Sector." *Quarterly Journal of Economics* 112(3):663-91.
- . 1998. "Private and Public Supply of Liquidity." *Journal of Political Economy* 106(1):1-40.
- IFIAC (International Financial Institution Advisory Commission). 2000. "Report of the International Financial Institution Advisory Commission." Allan H. Meltzer, Chair, Washington, DC.
- Jacklin, C., and S. Bhattacharya. 1988. "Distinguishing Panics and Information-Based Bank Runs: Welfare and Policy Implications." *Journal of Political Economy* 96(3):568-92.
- Jeanne, O. 2000. "Foreign Currency Debt and the Global Financial Architecture." *European Economic Review* 44(4-6):719-27.
- Jeanne, O., and C. Wyplosz. 2001. "The International Lender of Last Resort: How Large is Large Enough?" CEPR Discussion Paper 2842. Centre for Economic Policy Research, London.
- Kaminsky, G., and C. Reinhart. 1999. "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems." *American Economic Review* 89(3):473-500.
- Kashyap, A. K., R. Rajan, and J. Stein. 2002. "Banks as Liquidity Providers: An Explanation for the Coexistence of Lending and Deposit-Taking." *The Journal of Finance* 57(1):33-73.
- La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. Vishny. 1998. "Law and Finance." *Journal of Political Economy* 106(6):1113-55.
- Litan, R. 1987. *What Should Banks Do?* Washington, D.C.: Brookings Institution Press.
- Matutes, C., and X. Vives. 2000. "Imperfect Competition, Risk Taking and Regulation in Banking." *European Economic Review* 44(1):1-34.
- Mishkin, F. 1999a. "Financial Consolidation: Dangers and Opportunities." *Journal of Banking and Finance* 23(2-4):675-91.
- . 1999b. "Lessons from the Tequila Crisis." *Journal of Banking and Finance* 23(10):1521-33.
- Moreno-Villalaz, J. L. 1999. "Lessons from the Monetary Experience of Panama: A Dollar Economy with Financial Integration." *CATO Journal* 18(3):421-39.
- Morris, S., and H. S. Shin. 2000. "Rethinking Multiple Equilibria in Macroeconomic Modelling." Cowles Foundation Discussion Papers 1260. Cowles Foundation, Yale University. New Haven.
- Nakamura, L. 1988. "Loan Workouts and Commercial Bank Information: Why Banks Are Special." Working Paper 89-11. Federal Reserve Bank of Philadelphia.
- Petersen, M., and R. Rajan. 1994. "The Benefits of Lending Relationships: Evidence from Small Business Data." *Journal of Finance* 49(1):3-37.
- Postlewaite, A., and X. Vives. 1987. "Bank Runs as an Equilibrium Phenomenon." *Journal of Political Economy* 95(3):485-91.

- Powell, A. 2001. "A Capital Accord for Emerging Economies?" Policy Research Working Paper 2808. World Bank, Washington, D.C.
- Pulley, L. B., and D. B. Humphrey. 1993. "The Role of Fixed Costs and Cost Complementaries in Determining Scope Economies and the Cost of Narrow Banking Proposals." *Journal of Business* 66(3):437–62.
- Radelet, S., and J. D. Sachs. 1998. "The East Asian Financial Crisis: Diagnosis, Remedies, Prospects." *Brookings Papers on Economic Activity* 28(1):1–74.
- Rochet, J.C., and X. Vives. 2004. "Coordination Failures and the Lender of Last Resort: Was Bagehot Right After All?" *Journal of the European Economic Association* 2(6):1116–47.
- Rodrik, D., and A. Velasco. 1999. "Short-Term Capital Flows". NBER Working Paper 7364. National Bureau of Economic Research. Cambridge, Mass.
- Sachs, J., A. Tornell, and A. Velasco. 1996. "Financial Crises in Emerging Markets: The Lessons from 1995." NBER Working Paper 5576. National Bureau of Economic Research, Cambridge, Mass.
- Tirole, J. 2002. *Financial Crises, Liquidity, and the International Monetary System*. Princeton, N.J.: Princeton University Press.
- Vives, X. 2001. "Restructuring Financial Regulation in the European Monetary Union." *Journal of Financial Services Research* 19(1):57–82.
- 2002. "External Discipline and Financial Stability." *European Economic Review* 46(4–5):821–8.
- Wallace, N. 1996. "Narrow Banking Meets the Diamond–Dybvig Model." *Federal Reserve Bank of Minneapolis Quarterly Review* 20(1):3–13.
- World Bank. 1998. "Argentina: Financial Sector Review." Sector Report 17864. Washington, D.C.

---

# Access to Financial Services: A Review of the Issues and Public Policy Objectives

---

Stijn Claessens

---

*This article reviews the evidence on the importance of finance for economic well-being. It provides data on the use of basic financial services by households and firms across a sample of countries, assesses the desirability of universal access, and provides an overview of the macro-economic, legal, and regulatory obstacles to access. Despite the benefits of finance, the data show that use of financial services is far from universal in many countries, especially developing countries. Universal access to financial services has not been a public policy objective in most countries and would likely be difficult to achieve. Countries can, however, facilitate access to financial services by strengthening institutional infrastructure, liberalizing markets and facilitating greater competition, and encouraging innovative use of know-how and technology. Government interventions to directly broaden access to finance, however, are costly and fraught with risks, among others the risk of missing the targeted groups. The article concludes with recommendations for global actions aimed at improving data on access and use and suggestions on areas of further analysis to identify constraints to broadening access.*

Finance matters for economic development. There is considerable evidence today for a strong causal relationship between the depth of the financial system (as measured, for example, by the supply of private credit or stock market capitalization) on the one hand and investment, growth, poverty, total factor productivity, and similar indicators on the other hand. Indeed, many empirical cross-country tests have shown initial financial development to be one of the few robust determinants of a country's subsequent growth. Finance also matters for the well-being of people beyond overall economic growth. Finance can help individuals smooth their income, insure against risks, and broaden investment opportunities. Finance can be particularly important for the poor. Recent evidence has shown that a more developed financial system can reduce poverty and income inequality.

Much of this evidence has focused attention on the importance of overall financial development. Yet banking systems and capital markets, especially in developing

countries, are often skewed toward those who are already better-off, catering mainly to large enterprises and wealthier individuals. Many segments of the enterprise and household sectors lack access to finance, likely impeding their growth and reducing their welfare. What are the barriers to wider access to financial services? Should broader availability of financial services be a public goal, and if so what are the best means of achieving it?

This article reviews the evidence on the importance of financial development for economic well-being; examines the concepts of access and use of financial services; provides data on the extent of use for a sample of countries; assesses the desirability of universal access; considers the macroeconomic, legal, and regulatory obstacles to access; and reviews the risks and costs associated with attempts to broaden the provision of access to finance. The article is structured around the following questions: Why the recent attention on access? What does access to finance mean? What evidence is there on access, and who has access and who does not? What are the constraints to access, and what can governments do to improve access? And what are possible international actions to improve access?

## Importance of Finance for Development

Financial development has received increased attention lately and has become a more important part of the development agenda, for several reasons. Evidence that financial development matters for growth has been accumulating over the last decade. Based on changes in economies and economic production, finance may have moved up in the ranking of barriers to growth. And there is an increasing perception that the distribution of finance has been skewed for households and enterprises. Each of these explanations is reviewed briefly here.

### *Evidence on Finance and Growth*

There is much more evidence today that finance contributes to growth. The empirical evidence is robust and available at the country, sector, and individual firm and household levels using various statistical techniques. Financial deepening has been shown to "cause" growth (Demirgüç-Kunt and Maksimovic 1998; Rajan and Zingales 1998; Beck, Levine, and Loayza 2000; for a review of the evidence see Levine 2005). A doubling of private sector credit to GDP is associated with a 2 percentage point increase in the rate of GDP growth (WorldBank 2001).<sup>1</sup>

Finance influences growth through many channels. Finance helps growth by raising and pooling funds, allowing more and more risky investments to be undertaken; by allocating resources to their most productive use; by monitoring the use of funds; and by providing instruments for risk mitigation. It is less the form in which these

services come—whether from banks or capital market—than that they are being provided efficiently—by a proper institutional and competitive environment—that matters for growth (Demirgüç-Kunt and Levine 2001; see also World Bank 2001). As such, it is difficult to assert that particular types of financial systems are more or less conducive to growth or that one type of system is more or less conducive to facilitating universal access to financial services.

Finance also helps to improve income distribution and poverty reduction through several channels. Foremost, finance helps through economic growth, thus raising overall income levels. Finance can help more specifically by distributing opportunities more fairly. There is evidence, although more recent, that finance matters especially for poor households and smaller firms. Controlling for reverse causality, Beck, Demirgüç-Kunt, and Levine (2004) find in cross-country studies on the link between finance and changes in inequality and poverty that financial development causes less income inequality. Clarke, Xu, and Zou (2003) also find that the level of inequality decreases as finance develops, and since the more concentrated income is the higher poverty is, finance thus helps reduce poverty.

Honohan (2004) shows that financial depth explains the level of poverty (number of people with incomes of less than \$1 or 2 a day). But he also finds that across countries the degree of microfinance penetration, often thought to be specifically useful for the poor, has no special effects on poverty. (Barr 2005 reviews the more general links between microfinance and financial development.) Other evidence, however, such as Morduch and Hayley (2002), finds some specific impact of microfinance on poverty. Microfinance has been found to reduce poverty by alleviating credit constraints, thus reducing child labor and increasing education, and by insuring against shocks (Morduch forthcoming). More generally, with a few exceptions, it is arguable that direct access of poor people to financial services can strongly affect the attainability of the Millennium Development Goals.<sup>2</sup> Even the goals that chiefly require upgrading public services in health and education also require that poor households be able to afford these services (Littlefield, Morduch, and Hashemi 2003).

### *Rising Importance of Finance as Economies Change*

As economic production changes and countries liberalize their real economies, it has become clearer that the degree of financial development strongly influences the ability of countries, firms, and individuals to make use of new growth opportunities. Finance matters for firms' growth opportunities, especially for small- and medium-size enterprises. Beck, Demirgüç-Kunt, and Levine (2005) show that while successful economies typically have large-, small-, and medium-size enterprise sectors, these sectors do not "cause" growth, alleviate poverty, or decrease income inequality.

Rather, it is the overall business environment—ease of firm entry and exit, sound property rights, and proper contract enforcement—that influences economic

growth. Finance, however, accelerates growth by removing constraints that impede small firms more than large firms.<sup>3</sup> Finance allows firms to operate on a larger scale, encourages more efficient asset allocation, and eases the entry of new firms (Klapper, Laeven, and Rajan 2004). Financial—and institutional—development thus helps to level the playing field for firms and countries, especially important in a global economy with rapidly changing growth opportunities.

### *Skewed Distribution of Finance*

While financial development in general is beneficial for growth and poverty, finance may not be available on an equal basis. Although hard to prove for a large sample of countries, there is increasing evidence that finance often benefits the privileged few, especially in developing countries. In normal times this has meant that finance is allocated on the basis of connections and nonmarket criteria, acting as an entry barrier (Rajan and Zingales 2003). In times of crises this has meant that the costs of financial crises are allocated unevenly, with the brunt borne by the poor. Halac and Schnukler (2003) show that financial transfers during crises are large and regressive and expected to increase income inequality. (See also Claessens and Perotti 2005 and references therein for more discussion of the uneven distribution of finance and the impact of financial reform on inequality.)

## What Does Access to Financial Services Mean, and How Do Access and Use Differ?

Access to finance is not the same as use of financial services. Access refers to the availability of a supply of reasonable quality financial services at reasonable costs, where reasonable quality and reasonable cost have to be defined relative to some objective standard, with costs reflecting all pecuniary and nonpecuniary costs. Use refers to the actual consumption of financial services. The difference between access and use can be analyzed in a standard demand–supply framework. Access refers to supply, whereas use is the intersection of the supply and demand schedules. Figure 1 shows the categories of use and access on a continuum (in reality some of the categories will overlap). Group A has access and use of financial services. Group B has access but does not want to use financial services (voluntary exclusion). Group C has no access and thus does not use financial services (involuntary exclusion).<sup>4</sup>

Access is thus equal to  $A + B$ . Those who use financial services (A) clearly have access. Zero use or voluntary exclusion (B) does not necessarily reflect unavailability of services nor does it necessarily mean rationing. The demand and supply schedules may be such that some households or firms have access to financial services but

**Figure 1.** Difference between Access and Use

A	B		C		
Current consumers of financial services	Voluntary exclusion		Involuntary exclusion		
	No need No awareness?	Assumed rejection inability to use due to price/income	Rejected: High risk / bad credit = No access	Rejected: Discrimination= No access	Excluded due to price, product, income, or respondent features=No access
	B1	B2	C1	C2	C3
	Population				

Source: Author's analysis.

decide not to use them because they have no need, have no savings, rely on **nonfi**-  
nancial means of transacting (barter), or decide the prices are too **high**.

Whether demand and supply intersect will depend on the relative costs of provid-  
ing financial services and the income of consumers. If the relative prices of financial  
services go down compared with the prices of other goods, some of those who volun-  
tarily excluded themselves may start to demand financial services. Availability of  
services is a necessary, but not sufficient, condition for use. The supply and demand  
schedules may fail to intersect, in which case there will be lack of access, so that  
some households or firms are involuntarily excluded (C). They may lack access  
because, for example, barriers to access the formal financial system are too high or  
costs are unreasonably high or because they do not have a credit record.

That use will vary from access is a standard demand and supply result and is well  
accepted. However, analytical financial research, beginning with Stiglitz and Weiss  
(1981), has shown that, given information asymmetries, lenders will adjust not  
only price (interest rates) but also quantity and because of adverse selection and  
moral hazard concerns may not be willing to provide any financing to some individ-  
uals or firms. Depending on the distribution of borrowers' risk and return and other  
fundamental factors, such as income levels and net worth, the supply curve can be  
backward bending, leading to quantity rationing. Such rationing means involun-  
tary exclusion on the consumer side but is a rational market response on the supply  
side. Determining empirically whether an individual or firm has access to financial  
services but chose not to use them or was rationed out is complex. The effects of  
adverse selection and moral hazard, for example, are very hard to separate empiri-  
cally (Karlan and Zinman 2005).

In practice, the borderlines between the three groups are even less precise. Use will vary more from access when there are nonprice barriers. Some individuals will not have access to financing because there are no distribution points of financial institutions in their area—the supply curve is vertical at zero for them. Nonprice barriers can interact with the prices charged for financial services. The costs of financing rise for customers whose credit history is not well known, deterring them from seeking financing or rationing them out of the market. But their lack of a credit history may arise from such barriers as a weak institutional environment, including poorly functioning credit information bureaus. Lack of access because banks do not serve a particular area or charge too much may arise because of a low level of competitiveness in the banking system.

Distinguishing use and access also depends on the aspect of finance being considered—savings mobilization, allocation of loanable funds (credit), payment facilitation, and insurance (see Bodie and Merton 1995 for a review of the functions of finance). For example, some individuals may have access to payments services but not to credit. For measurement purposes it is often hard to distinguish between these functions, as say an account at a bank provides both payment and insurance services and may also be the starting point of credit. This further complicates the access analysis.

Some analysts have tried to provide more specific definitions of access to financial services by categorizing the different dimensions to access. First is the dimension of availability: are financial services available, and if so in what quantity? Second is the question of cost: at what total price are financial services available, including the opportunity costs of having to wait in line for a teller or having to travel a long distance to a bank branch? Third is the range, type, and quality of financial services being offered. Following Morduch (1999), these dimensions can be identified as reliability, is finance available when needed; convenience, is access easy; continuity, can finance be accessed repeatedly; and flexibility, is the product tailored to individual needs.

Variants of these dimensions are used in other studies.<sup>6</sup> The point is that there are various dimensions to access, and consequently various dimensions in which access may be deficient. There can be deficient access geographically to branches and outlets; or deficient access socioeconomically. Or access can be deficient in an opportunity sense: the deserving do not have access. All of this makes it (even) more difficult to establish conceptually the degree of access, let alone to measure it.

## What Do Data About Use Tell Us?

These analytical questions on access and use indicate the difficulty of defining access. Empirically, documenting access faces the further challenge of limited data

on the degree of use of financial services. Although there is much data on financial sector development, there is very limited data on use of financial services, both for households and for firms, across countries (Emerging Market Economics 2005; Honohan 2005). Data are insufficient in all respects, making judgments on the causes of lack of access more difficult.

For a reasonable number of countries there are data from providers on households' use of basic financial services, such as the number of people with a bank account. These data are often obtained using commercial bank and central bank data or from surveys. Recently, data have been collected on the spread of microfinance services (CGAP 2004) using data from individual microfinance institutions (as collected by the microfinance information exchange). These cover the number of people with access to a savings account. Similarly, Beck, Demirgüç-Kunt, and Martinez Peria (2005) have compiled data from regulators for a sample of countries on the number of accounts and average loan and deposit size at commercial banks.

For some countries there are micro-based data from household surveys, such as the Living Standard Measurement Study coordinated by the World Bank. Some 27 of these have covered some dimensions of households' use of financial services (Gasparini and others 2005). Still, and with the exception of some industrial countries such as the Netherlands and Sweden, much of the data collected in these general household surveys is very basic and limited in the various dimensions of use and access (quantity, costs, and quality). Use of and access to credit have been difficult to document at the household level. Many countries do not even have data on the aggregate level of consumer credit, in part because nonbank financial institutions as well as banks provide credit.

Data on firms' use of and access to financial services are somewhat less limited. Considerable information is available on listed firms' financial structures and their use of external financing. Much less information is available on unlisted firms, especially on the financial structure of small firms' finance. Most data on smaller firms come from surveys, such as those conducted by the World Bank (World Bank Economic Survey and Investment Climate Assessments) or by national agencies such as the U.S. Federal Reserve Boards and the U.K. Bank of England. Some data come from central bank statistics (Central Bank of the East African States, BCEAO, for example, collects data on use) and advocacy groups (U.S. Small Business Administration, chambers of commerce, and equivalents). Again, the data are basic and limited in dimensions of use (quantity, costs, quality). Use of credit dominates data collection efforts, with use of savings services less of an issue, although payment services are important as well for firms. Furthermore, most data are collected on use of banking services, and much less information is available on the use of other financial services, such as insurance, leasing, factoring, and the like.

Although weak and often not comparable, available data show that use by households of banking services varies greatly. Many households in developing countries

do not have a bank account. With the main caveat that data are not easily comparable across countries and some of the numbers are only rough estimates, table 1 provides data on the degree to which households use a basic financial service provided by a formal financial institution (have a checking or savings bank account) across many countries. It shows that in most Organisation for Economic Co-operation and Development countries use is nearly universal, averaging 90 percent; in developing countries use is much lower, averaging 26 percent. The highest use of financial services from formal financial institutions is 59 percent in Jamaica. High use rates in some other countries may not be representative of the whole country as they apply to the population of the capital city only (Mexico) or to specific cities or regions (China, Colombia, and India) or urban areas (Brazil). For most of the other developing countries use of a basic bank account does not exceed 30 percent, and in the lowest income countries use is less than 10 percent of households.

Individuals obtain financial services through other means, including nonfinancial institutions (table 2). The microfinance information exchange data also show that financial services outside the banking system are often used. Thus, these numbers underestimate the degree of access to financial services, but they do show the large differences between industrial and developing countries in use of financial services from formal financial institutions.

The next question then is who are the unbanked households, and how do they differ between industrial and developing countries? Only revealed use and not access is observed. Thus, scenarios of zero transactions in which there is demand cannot be distinguished from those where there is lack of demand, although household and firm surveys provide some insight into the reasons behind the (lack of) demand. To the extent known, the profiles of the unbanked are as expected. Socio-economic characteristics such as income, wealth, and education play the largest roles in explaining observed use. Financial exclusion is often part of a broader pattern of exclusion that includes education and jobs and other areas of life. Households that use credit have a different profile from those that have bank and savings accounts, and the profile is affected by income and wealth characteristics, as it tends to be the richer who borrow.

A comparison between the United States and Latin American countries shows some similarities between otherwise very different countries in which people do not want to bank (table 3). After banks barriers, convenience, trust, and savings are important considerations for households that do not seek financial services from banks in all countries except Colombia. Macroeconomic factors can play an important role in demand, as when banking and financial crises have undermined the confidence of the public in the formal financial system. Colombia, for example, has had few banking crises, and the percentage of unbanked who cite mistrust as a reason not to use financial services is much lower than in the other two Latin American countries, which have had more crises.

**Table 1.** Share of Households with Access to a Bank Account or Using Financial Services

Country	Source	Date of survey	Number of households	Share of household (percent)					
				That saved money in the past 12 months	That used formal financial institutions to save	That used informal finance to save	That borrowed money in the past 12 months	That used formal financial institutions to borrow	That used informal finance to borrow
<i>Developing group</i>									
Armenia	LSMS	1996	4,920	17.1	8.9	0.2	—		4.5
Bosnia and Herzegovina	LSMS	2001	5,400	—	—	—	21.8	6.2	15.8
Botswana	FINSCOPE	2003	530		47.0	25.7	—	11.7	29.1
Brazil (11 urban areas)	SAPS	2002	2,000		42.7	45.5		—	—
Bulgaria	LSMS	2001	2,633	—	—	—	5.4	5.4	
China (Hebei and Liaoning)	LSMS	1995–97	787	82.5	41.9	13.3	28.1	5.2	24.7
Colombia (Bogota city)	—	—	—	—	41.2		—	—	—
Côte d'Ivoire	LSMS	1988	1,600	88.1	24.8	—	23.8	3.2	21.1
Ghana	LSMS	1998/99	5,998	12.0	—		39.1 <sup>a</sup>	3.3	32.0
Guatemala	LSMS	2000	7,276	18.1	17.8	0.4	31.8	23.5	7.4
Guyana	LSMS	1992/93	1,819	15.7	13.7	3.6	4.7	1.3	2.5
India	AIDIS	1991	57,031		—		26.9	11.8	19.0
India (Uttar Pradesh and Andhra Pradesh)	RFAS	2003	6,000	—	47.5 <sup>b</sup>				
Jamaica	LSMS	1997	2,020	68.1	59.4	17.8	10.5	1.9	5.9
Kenya	Estimate			—	10.0	—	—	—	—
Kyrgyz Republic	LSMS	1998	2,979	11.4	1.3	10.1	6.1	0.3	5.3
Lesotho	FINSCOPE	2003	534	—	17.0	11.0	—	—	6.0
Mexico (Mexico City)	—	—	—	—	25.0	—	—	—	—
Morocco	LSMS	1990/91	3,323	15.5			22.0	3.6	19.3
Namibia	FINSCOPE	2003	810	—	28.4	0.9	—	5.3	15.2

*(Continued)*

**Table 1.** (Continued)

Country	Source	Date of survey	Number of households	Share of household (percent)					
				That saved money in the past 12 months	That used formal financial institutions to save	That used informal finance to save	That borrowed money in the past 12 months	That used formal financial institutions to borrow	That used informal finance to borrow
Nepal	LSMS	1996	3,373	—	—	—	57.0	12.9	50.0
Nicaragua	LSMS	1998/99	4,209	6.5	4.7	0.2	22.5 <sup>c</sup>	7.6	7.8
Pakistan	LSMS	1991	4,800	23.6	12.2	14.5	30.3	1.1	29.4
Panama	LSMS	1997	4,945	—	—	—	1.5	0.8	0.8
Peru	LSMS	1994	3,623	25.2	—	—	16.6	2.1	14.5
Romania	LSMS	1994/95	2,4560 <sup>d</sup>	94.3	22.5	1.4	15.9	6.0	11.4
South Africa	LSMS	1993	9,000	—	—	—	44.8	4.9	42.6
South Africa	FINSCOPE	2004	2,988	—	46.0	—	—	—	—
Swaziland	FINSCOPE	2003	604	—	35.3	19.5	—	4.1	16.1
Tanzania	Estimate	—	—	—	5.0	—	—	—	—
Uganda	Estimate	—	—	—	< 5	—	—	—	—
Viet Nam	LSMS	1997/98	6,002	89.9	8.7	12.8	49.1	26.1	30.4
<i>Developed group</i>									
Austria	—	—	—	—	81.4	—	—	—	—
Belgium	—	—	—	—	92.7	—	—	—	—
Denmark	—	—	—	—	99.1	—	—	—	—
Finland	—	—	—	—	96.7	—	—	—	—
France	—	—	—	—	96.3	—	—	—	—
Germany	—	—	—	—	96.5	—	—	—	—
Greece	—	—	—	—	78.9	—	—	—	—
Ireland	—	—	—	—	79.6	—	—	—	—
Italy	—	—	—	—	70.4	—	—	—	—
Luxembourg	—	—	—	—	94.1	—	—	—	—

**Table 1.** (Continued)

Country	Source	Date of survey	Number of households	Share of household (percent)					
				That saved money in the past 12 months	That used formal financial institutions to save	That used informal finance to save	That borrowed money in the past 12 months	That used formal financial institutions to borrow	That used informal finance to borrow
Netherlands	—	—	—	—	98.9	—	—	—	—
Portugal	—	—	—	—	81.6	—	—	—	—
Spain	—	—	—	—	91.6	—	—	—	—
Sweden	—	—	—	—	98.0	—	—	—	—
United Kingdom	—	—	—	—	87.7	—	—	—	—
United States	SCF	2001	4,449	—	90.9 <sup>e</sup>	—	75.1 <sup>f</sup>	—	—

—, not available.

Note: Definitions of formal and informal financial institution vary greatly among countries due to differences in survey questionnaires. Generally speaking, for savings, formal financial service providers include banks (public or private), cooperatives, and credit unions. For a few countries other financial institutions, such as security firms and postal savings, are also included in formal. Informal includes others that provide financial services, except for savings at home. Microfinance institutions and nongovernmental organizations are included in informal, as are rotating savings and credit associations, tontines, moneylenders, pawnshops, *ususus*, and *stokvels*. For borrowing, the same definitions are followed, with person to person borrowing included in informal.

<sup>a</sup>Households who paid off rents.

<sup>b</sup>Deposit accounts only.

<sup>c</sup>Credit purchases.

<sup>d</sup>Based on number of households in the household roster files.

<sup>e</sup>Family holding some type of transaction account—a category comprising checkings, savings, money market deposit accounts, money market mutual funds, and call accounts at brokerages.

<sup>f</sup>percentage of family holding any debt.

Sources: The main sources are Living Standard Measurement Study (LSMS) surveys, with household responses averaged for each country, and Napier (2005) for many Southern African countries. For the EU countries, Pesaresi and Pilley (2003). For the United States, Board of Governors, U.S. Federal Reserve System (2004). For Brazil, Colombia, India, and Mexico, Kumar and others (2004), Basu and Srivastava (2005), and Caskey, Solo, and Durán (2004), except that early data for India are from the regular Indian household surveys. For Kenya, Tanzania, and Uganda, Peachey and Roe (2004).

**Table 2.** Distribution of Savings Deposits in Four Countries (percent of total)

	<i>Brazil</i>	<i>India</i>	<i>Colombia</i>	<i>Mexico</i>
Banks	95 (54 private; 41 public)	90 (30 rural regional banks)	85	96
Cooperatives	0	7	14	
Post office	—	2	—	
Family and friends	4	—	—	
Others	1	1	1	4

—, not available.

*Note:* Response to question: What other savings and deposit facilities are being used?

*Source:* Kumar and others 2004.

**Table 3.** Reasons the Unbanked Do Not Use Banks: A Comparison of Five Countries (percent of total)

	<i>United States</i>	<i>Mexico</i>	<i>Colombia</i>	<i>Brazil</i>	<i>India</i>
<i>Demand limitations</i>					
No need, no savings	53	7	16		75
No awareness					18
Supply limitations (bank barriers: high costs, minimum balances, documentation)	45	70	78	42	
<i>Perceptions of service</i>					
Safety, mistrust	18	16	3	25	
Lack of documentation	10	3			
Privacy	22	2			
Inconvenience (location and hours)	10	2			
Other reasons			3	33	

—, not available.

*Source:* Kumar and others 2004.

Unbanked households in the United States and Mexico, two countries at different levels of development, also display very similar characteristics, with the exception of home ownership (table 4). The costs of being unbanked vary considerably, however, as alternatives are much fewer and more costly in Mexico. In the lowest income segment the costs of being unbanked are estimated at 2.5 percent of median income in the United States and 5 percent in Mexico (Caskey, Duran, and Solo 2004; see also Solo 2005).

Although weak and often not comparable, some data on firm's access to financing have more recently become available from the World Bank Investment Climate Assessments that have been conducted in the last few years. About a quarter of the firms on average complain that lack of access to external financing is a main or severe obstacle to the operation or growth of their business (table 5). There are large variations; from less than 7 percent for Latvia and Lithuania to more than 50 percent

**Table 4.** Who are the Unbanked? Comparison of the United States and Mexico (percent)

	<i>United States</i>	<i>Mexico</i>
<i>Similarities</i>		
Below median income	79	90
Less than high school	56	51
Marginalized in socioeconomic terms	90 <sup>a</sup>	60 <sup>b</sup>
<i>Differences</i>		
Percentage of total	9.1	75 <sup>c</sup>
Home ownership	7.8	63 <sup>c</sup>

<sup>a</sup>Latino and African American.

<sup>b</sup>Informal sector.

<sup>c</sup>In Mexico City.

*Source:* Solo, Caskey, and Duran 2004.

for several countries and a high of 60 percent for Brazil. Of course, these raw scores on firms' complaints about financing availability cannot be taken as an indicator of lack of access. They are heavily affected, for example, by short-term conditions in financial markets and macroeconomic policies, as shown by the comparison between Estonia, where real interest rates are in the low single digits, and Brazil, where real interest rates are more than 10 percent.

Somewhat similar to the question of unbanked households is that about unbanked firms. To the extent that we know, profiles are as expected, with the size of the firm (and, related, its age) especially important. Table 5 suggests this, as the share of large firms with complaints is less than the share of the smallest firms—on average some 8 percentage points difference but sometimes as much as 10–20 percentage points. Across a large sample of countries and controlling for other factors, it has also been found that size has the strongest effects on access to credit (Beck, Demirgüç-Kunt, and Maksimovic 2005; see also Beck and others 2005).<sup>7</sup> For Brazil size was found to be more important than performance and other variables, suggesting quantitative limitations to credit access (Francisco and Kumar 2005). The impact of size on credit was found to be greater for long-term loans in Brazil and in many other countries.

However, size may reflect not only profitability and financial and legal collateral but also political collateral. This is particularly so in developing countries, where lending is often on the basis of relationships and connections, which are often political. In countries with well-developed financial systems, size constraints can be overcome. Many banks in industrial countries lend to small single proprietor firms, sometimes without requiring collateral, financial statements, or other requirements. Thanks to the spread of technological advances such as automated credit scoring, and banks in developing countries are also becoming active in these forms of financing.

**Table 5. Complaints by Firms about Lack of Access to External Financing (percentage of firms)**

<i>Country</i>	<i>Year</i>	<i>Country average</i>	<i>Small (1–49 employees)</i>	<i>Medium (50–249 employees)</i>	<i>Large (250+ employees)</i>
Albania	2002	14.3	16.0	14.8	0.0
Algeria	2002	53.1	55.8	45.4	44.0
Armenia	2002	21.8	23.0	19.1	18.2
Azerbaijan	2002	13.0	12.6	24.0	4.2
Bangladesh	2002	41.6	42.5	45.8	36.4
Belarus	2002	25.8	30.7	12.5	18.9
Bosnia and Herzegovina	2002	22.5	23.2	19.4	25.0
Brazil	2003	60.5	61.9	60.6	51.9
Bulgaria	2002	38.5	45.2	18.4	29.7
Cambodia	2003	9.4	9.3	5.3	12.2
China	2002	22.8	16.9	20.9	30.2
Croatia	2002	24.9	26.9	24.2	19.2
Czech Republic	2002	25.1	28.8	25.6	7.0
Ecuador	2003	44.9	49.0	36.0	50.0
El Salvador	2003	31.0	36.5	19.6	19.5
Eritrea	2002	52.3	55.0	52.6	33.3
Estonia	2002	12.1	11.8	16.7	9.1
Ethiopia	2002	42.8	43.5	48.0	27.9
Georgia	2002	10.5	10.1	15.4	5.9
Guatemala	2003	34.1	36.4	33.3	22.0
Honduras	2003	50.8	56.2	51.4	20.7
Hungary	2002	21.6	26.1	16.7	9.5
India	2002	18.3	—	—	—
Indonesia	2003	17.5	16.0	16.6	19.5
Kazakhstan	2002	11.7	9.9	11.4	20.6
Kenya	2003	44.1	60.5	30.6	39.3
Kyrgyz Republic	2002	16.0	15.1	20.5	5.9
Latvia	2002	6.1	5.2	4.2	12.0
Lithuania	2002	6.7	5.7	12.1	4.6
Macedonia, FYR	2002	13.3	13.2	8.7	22.2
Moldova	2002	25.2	27.1	15.2	30.0
Nicaragua	2003	54.4	57.2	47.2	12.5
Pakistan	2002	37.6	38.6	34.6	27.5
Peru	2002	50.2	50.7	62.5	66.7
Philippines	2003	13.5	15.0	14.4	8.9
Poland	2002	32.7	36.5	22.6	32.2
Romania	2002	29.7	32.9	25.4	24.3
Russian Federation	2002	20.3	21.0	18.0	20.7
Serbia and Montenegro	2001	33.6	30.8	33.3	42.9
Slovak Republic	2002	29.6	32.7	27.3	20.8
Slovenia	2002	8.2	9.9	4.0	0.0
Tajikistan	2002	22.5	24.5	15.6	25.0
Tanzania	2003	48.3	54.8	38.2	27.3

**Table 5.** (Continued)

<i>Country</i>	<i>Year</i>	<i>Country average</i>	<i>Small (1–49 employees)</i>	<i>Medium (50–249 employees)</i>	<i>Large (250+ employees)</i>
Turkey	2002	17.3	16.1	25.7	9.8
Uganda	2003	45.0	47.3	39.0	33.3
Ukraine	2002	26.4	31.1	16.3	18.2
Uzbekistan	2002	26.5	26.0	21.4	35.3
Zambia	2002	53.7	65.2	51.6	37.1
Percentage of firms that say access to financing presents main or severe obstacles to the operation and growth of their business		28.78	31.06	26.43	23.22
		48	47	47	47

— not available.

*Note:* Percentage of firms that say access to financing presents main or severe obstacles to the operation and growth of their business.

*Source:* World **Bank** Investment Climate Assessments ([www.rru.worldbank.org](http://www.rru.worldbank.org)) downloaded on February 4, 2005.

And in the most developed financial markets, universal access to basic financial services, including some forms of credit, is essentially ensured for households.

That use is not universal may reflect lack of demand rather than lack of access: many households and firms may not use financial services, despite having access to some financial services. But with use so low in many countries, the question naturally arises whether this is because the supply of financial services is limited. And if supply is limited, is it because financial service providers consider some households and firms as less attractive customers and are therefore unwilling to extend financial services? Or is it because there are barriers to supply? If there are barriers, the policy question is whether these can be removed without creating other economic costs or risks. If the lack of supply is due to some market failure, does there still remain a need for government intervention?

## Institutional Barriers to Access

Institutional or supply barriers to access include specific constraints of financial institutions and barriers arising from the overall institutional environment, which can include a weak legal system, weak information infrastructure, and lack of competitiveness in the banking system. In the terminology of Beck and de la Torre (2005), options to expand supply would thus include moving toward the country's

access possibilities frontier through individual financial institution solutions and expanding the country's access possibilities frontier through country actions.

### *Individual Financial Institution Constraints*

Households and firms often state that they do not use financial services because the services are too costly or not the right type. Households often mention problems of high minimum deposits and high administrative burdens and fees. Many small borrowers are deterred by the high fixed costs of applying for loans and the often-high rejection rates. Financial institutions may demand collateral, which poor borrowers typically lack. Formal financial services may also entail nonpecuniary barriers, such as high literacy requirements.

Households and firms may instead seek financial services from informal sources. Individuals needing funds for investment may rely on family and friends. People wanting to transmit payments, whether domestic or international, may rely on informal networks, although at higher costs. This is most obvious in the transmission of international remittances, where unit costs of informal mechanisms can be very high. To wire \$100 from New York to Mexico costs \$9 for the banked and \$19 for the unbanked, plus an unknown exchange rate spread in both cases (Casley, Duran, and Solo 2004). Yet these informal mechanisms are often preferred because of other, nonpecuniary barriers.

When the environment is sufficiently competitive, financial institutions can be expected to adjust product features and costs as much as possible, given their cost structures. Yet financial service providers commonly respond that they do not serve poor households and small firms because the risk and costs are too high. Financial institutions do not find it profitable enough to offer appropriate financial services to some segments.

There may be a variety of reasons for the lack of provision of appropriate products and services. Banks may have problems offering financial services to all households. It may be too costly to provide the physical infrastructure in areas of low population density or where there is a lack of security. High transaction costs for small volumes are often mentioned as constraining financial service providers from broadening access. Small borrowers borrow frequently, for example, and repay in small installments. They consequently do not want financial products with high per unit costs, yet for banks costs are often similar regardless of transaction size. Households and firms in developing countries may seek financing or insurance for specific purposes (important life events such as marriage, healthcare, or specific crop insurance) for which contracts are difficult to design. Firms may be underserved for the same reasons. Small firms seek different products than large enterprises, such as payment services for small amounts, and banks may not consider these firms attractive as clients. Small markets may make it more difficult to develop or roll out new products specifically useful for these markets.

The fixed costs in financial intermediation thus make providing services for small clients, by small institutions, and in small markets hard. At the same time economies of scale lead to decreasing unit costs as transaction volumes increase, making some specialization attractive. Although better cost management can lower unit costs, there are limits to cost management at the level of an individual institution, as evidence on the economies of scale for banks in mature financial markets shows (see Berger and Humphrey 1997 for a survey). Evidence on microfinance institutions also reveals economies of scale (Honohan 2004). The proliferation of microfinance institutions in many countries has not necessarily benefited final clients as much as possible, because few institutions have reached the scale necessary for efficient financial services provision. Similar constraints arise at the country level, where many financial systems are very small (less than a few billion dollars equivalent, smaller than a very small bank in most industrial countries), hindering effective financial services provision (Hanson, Honohan, and Majnoni 2003). Scale for effective financial services provision may not exist in all countries, at least not using traditional, local financial services providers.

Banks and other financial institutions can move closer to the "access frontier," however, through innovation. Sometimes prodded by government and public opinion, they can make their products more suited to low-income households. In South Africa in 2004 the country's principal banks launched a low-cost bank account aimed at extending banking services to the black majority. The country's four big retail banks along with the post office's Postbank launched the *Mzansi* account. Set up under a financial sector charter agreed on by the industry in 2003, the account requires a minimum deposit of 20 rands (about \$4) and is aimed at providing access to financial services to some 13 million low-income South Africans without prior access to bank accounts. Whether this will be profitable and sustainable is to be seen, but the initial take-up has been promising (Napier 2005). The sharp drop in the costs of international remittances (Orozco 2004, Maimbo and Ratha 2005) also suggests that there is still room for moving closer to the frontier. De la Torre, Gozzi, and Schmukler (2005) provide other examples of innovative approaches for enhancing access for small firms.

Some of these recent innovative and low-cost solutions in delivering financial services suggest that the limits to adapting products to the needs of a broader class of borrowers have not yet been reached. For many of the mismatches between potential demand and supply, it is thus not clear whether there is a market failure and if so what the source is. Time will tell whether financial institutions will offer the right products, properly priced, and whether financial institutions operating at the right scale and with the right technology will enter certain markets. The fact that they do not yet do so may mean that it is not profitable, given the institutional environment they face in a particular market and given current technology.

## *Institutional Environment Constraints*

Although there is much analysis of what affects financial sector development and what role the institutional environment plays (World Bank 2001), evidence on what affects households' and firms' access to financial services is very limited. What evidence exists though gives some insights on the most binding constraints.

Across countries it is clear that there are some elements of overall development, including greater use of advanced technology, that allow banks in more industrial countries to offer financial services profitably to lower income segments (Beck, Demirgüç-Kunt, and Martinez Peria 2005). Of course, the incomes of the lower income segments in these more industrial countries are higher than the incomes of the lower income segments in developing countries, so it does not mean that the same technology can also reach the lower income segments in developing countries.

For microfinance it appears that access for the poor or the near-poor is worse in countries with higher GDP per capita, in countries with poorer institutional quality, and in countries with a smaller market (Honohan 2004). This suggests that good country institutions and a large potential market help the microfinance industry to grow. The same analysis shows that a poorer quality in the main banking system discourages the spread of microfinance institutions. Specifically, countries with higher spreads and higher profitability in their main banking system have fewer microfinance institutions. This suggests that more competition in the banking system can foster greater access to financial services, including those from microfinance institutions.

The use of savings and payments services also appears to be a function of distribution networks, including those of postal systems, saving banks, and other specialized financial institutions. In Brazil for example, the size and scope of branch networks, as well as the split between public and private banks and domestic and foreign banks, play a role in the degree of use of financial services (Kumar and others 2004; see also Kumar 2005). In other markets more specialized financial institutions such as savings banks and other proximity banks that have, besides profitability, the objective of providing financial services have broadened usage (Peachey and Roe 2004). For a sample of 91 countries Beck, Demirgüç-Kunt, and Martinez Peria (2005) show that countries with better developed financial systems and more efficient banks have wider branch and automated teller machine penetration and that the use of deposit and loan services is more evenly distributed among banking clients. These findings suggest that what is driving use is not purely the scope for profitable banking but also the overall institutional environment and level of development.

There has been more analysis of the access of small firms to financial services, and evidence suggests that the institutional environment matters even more than for households (see Berger and Udell 2005 for a review of the conceptual issues). This is

particularly so on the credit side. The absence of credit information, difficulty in registering and recovering collateral, and problems with contract design and enforcement can make lending especially difficult. Credit services may consequently be limited to entrepreneurs with credit history, (political) connections, or immovable collateral, such as real estate. Even when a business is viable, there may be little reliance on past records or expected future performance. In many countries there are problems of uncertain repayment capacity arising from volatile income and expenditures. Especially, new and smaller firms often have high exposures to these systemic risks (for example, macroeconomic volatility, financial crises, defaults by governments, and arbitrary taxation).

There is empirical evidence on the importance of these barriers. The quality of legal systems, property rights, and mechanisms for reliable information have been found to be especially important for small firms (Beck and others, 2004; Beck, Demirgüç-Kunt, and Maksimovic 2005). Small firms and firms in countries with poor institutions use less external finance, especially bank finance. Better protection of property rights increases use of external finance by small firms significantly more than by large firms, mainly because of more bank and equity finance. It also appears that substitutes for bank finance are imperfect; for example, small firms do not use disproportionately more leasing or trade finance compared with larger firms.<sup>8</sup> Beck, Demirgüç-Kunt, and Martinez Peria (2005) find that firms in countries with higher levels of financial system development and greater outreach report lower financing obstacles, with the association stronger in less economically industrial countries. This impact of outreach on financing obstacles does not vary with the degree to which the banking system is government-owned—government-owned banks do not "solve" this access problem.

Analysis at the country level has been more limited, but it provides some insights into what may be driving use. Government interference can distort risk-return signals, making it hard for formal financial institutions to offer attractive products. Interest rate regulations can interfere with the abilities of financial service providers to offer saving or lending instruments profitably. Administrative regulations and procedures can create high transaction costs and barriers for dealing with formal financial institutions. Many countries have customer identification requirements, the so-called Know your customer rules, which limit the ability to offer simple banking products. The recent focus on antimoney laundering and counterterrorism financing has led to laws that can adversely affect the provision of financial services, as it has threatened to do in South Africa (Napier 2005).

In addition to hindering the activity of existing financial services providers, regulation can discourage the emergence of financial institutions more suited to the needs of lower-income households or smaller firms. Rigidity in chartering rules, high minimum capital adequacy requirements, restrictions on funding structures, excessive regulation and supervision, and overly strict accounting requirements and

other rules can prevent microfinance institutions and smaller financial institutions from emerging. In South Africa bank regulation and supervision were being extended to microfinance institutions, which reduced their capacity to offer financial services profitably to the lower-income segments of the populations (Glaessner and others 2004). Separate charters may be useful, with the required structures depending on such factors as whether the institution borrows, takes deposits, or is owned by its members (Christen, Lyman, and Rosenberg 2003).

With these and other regulatory and supervisory requirements, tradeoffs arise, however, as the requirements are meant to serve other public policy purposes, such as financial stability and integrity. There are also tradeoffs in facilitating the mainstreaming of microfinance institutions. Jansson, Rosales, and Westley (2004) argue, for example, that new institutional forms should not be created for microfinance institutions unless there are several mature and well-managed nonprofit organizations ready to transform into such financial intermediaries and the existing institutional forms—such as banks or finance companies—are unusable (due to high minimum capital requirements, for instance) or too limited because of operational restrictions (such as the inability to mobilize deposits).

There is consequently a need to evaluate the value of regulatory approaches from an overall welfare point of view. Although approaches have to strike the right balance, they can be adjusted to enhance the supply of financial services. In many countries, for example, antipredatory lending laws are needed rather than usury laws, which prevent small borrowers from getting access to credit at all, even at high interest rates. Also, simplifying truth in lending requirements for small-scale lending, rather than applying the extensive small-print type regulations many countries have, can be useful to facilitate the supply of financial services. Adapting regulations can also mean facilitating multiple forms of financial services provision. That may involve considering savings mobilization separately from credit extension. Many households are interested in savings and payment services only, not in credit services. These types of financial services provision may require different forms of regulation and supervision.

Finally, much regulation is aimed at protecting savers and borrowers against misuse and risks, yet it may not be effective in developing countries given the lack of supervisory capacity, independence, and effective checks and balances and may end up impeding access (Barth, Caprio, and Levine 2005). The general level of financial literacy may need to be increased, as is actively being done in some countries. Consideration also needs to be given to educating people on the risks of (new) financial services and different types of financial service providers, so that people can strike the right balance between risks and benefits.

Improvement in institutional infrastructure is an area where progress can clearly be made in furthering access in many developing countries. Better legal, information, payments, distribution, and other infrastructure are needed. Such work is

already under way by many governments, multilateral financial institutions, and others, but it will take time. Other policy steps can be useful to increase access. The evidence reviewed by Honohan (2004) suggests that an important way to enhance access is by improving competition in the banking system. This is often easier to do than improving the institutional environment.

Increased competition can be applied to all segments of the financial sector. Smaller and nonbank financial institutions can be allowed greater use of existing networks, for example. In many countries access to the payments system is limited to a club of large banks. Information sharing is restricted in many countries to incumbent banks and formal financial institutions. This together with the limited existence of (private) credit bureaus is making it difficult for other financial institutions to provide financial services (Miller 2001). Few countries, for example, allow nonbank financial institutions and entities such as department stores access to bank information, making it more difficult for them to provide financial services to low-income households.<sup>9</sup> Yet, lower-income people often get their credit from these non-financial institutions. In Mexico, for example, close to 50 percent of credit for those with no banking relationship comes from department stores (Caskey, Durán, and Solo, 2004).

Although some of these changes are technically easy to adopt, competition policy is complex, especially in small markets with little institutional capacity, and political economy factors can prevent progress. A credible competition agency is required, for which the institutional requirements are quite high. Unchecked competition may not be the first-best choice. Allowing any new party to open a credit bureau can undermine the incentive structure for entities to provide accurate information while requiring financial institutions to disclose all types of information can undermine their willingness to enter relationships with their clients out of fear that competitors will take away their business. Furthermore, even in industrial countries questions arise on how best to address the many networks that exist in financial services (payments system, credit bureaus, and distribution networks), which raise special competition policy issues. Answers here are not obvious.

In addition to the general view that competition can help with access, there is specific evidence that allowing greater entry by foreign banks can enhance access (Clarke and others 2003 review the evidence). A study on borrowers' perceptions across 36 countries found that reported financing obstacles were lower in countries with high levels of foreign bank penetration (Clarke, Cull, and Martinez Peria 2001). The same study found strong evidence that even small enterprises benefited and no evidence that they were harmed by the presence of foreign banks. The channels appear both competitive pressures of foreign banks on the domestic banking system, forcing local banks to lend to smaller firms, and direct provision of financial services by foreign banks. A Latin American study found that foreign banks with small local presence do not appear to lend much to small businesses but that large foreign banks

in many cases surpass large domestic banks in such lending (Sanchez and others 2005).

There are also many examples of the effects of foreign bank entry. In Mongolia, with an income per capita of less than \$500 and a very rural-based economy, the government-owned Agricultural Bank of Mongolia (Khan Bank) was placed in receivership in 1999 after many years of operating deficits, loan losses, and a failed attempt at privatization. In March 2003 HS Securities of Japan bought Khan Bank for \$6.85 million. Khan Bank now operates a network with 379 points of service throughout Mongolia, greater than any of the other 16 banks operating in the country (and up from 269 when new management took office). Today, one of two Mongolian households is reportedly a client of Khan Bank, and it seems to continue to expand its branch network and services (World Bank 2006).

In addition to the direct provision of financial services, foreign bank entry has indirect effects on the overall banking system, such as greater financial stability and improved efficiency of financial intermediation (Clarke and others 2003). These two effects can make the local banking environment more conducive to lending, including to lower-income segments, and can put pressures on local banks to engage more in lending to lower-income segments as profitability in other segments declines.

The impact on access of foreign competition in securities markets is less obvious. Globalization has meant that large firms have been accessing international financial markets. In some developing countries this has reduced domestic stock market liquidity, possibly hurting access to finance by smaller firms. At the same time relaxing the financing constraints of large firms through access to international markets can ease the financing constraints of small firms that benefit indirectly, such as through trade credit arrangements. On net it is not yet clear whether small firms lose or gain from globalization and increased competition in securities markets.

## Role of Specific Interventions and Technological Improvements

Recent country experiences have shown that specific interventions besides the removal of barriers and improvements in the institutional environment can enhance access. In India, for example, discussions are under way to use existing networks (for example, the postal system) for the delivery of new financial services by other public and private providers. The idea is that use of the technology and information backbone of existing public or other networks need not be limited to one provider. Many countries have large networks of post offices that could be used to allow various financial institutions to offer electronic finance services. In Brazil where the post office has a presence in 1,738 of the more than 5,000 municipalities without a bank outlet, the government auctioned the exclusive right to distribute financial services

through the post offices in 2001. Although this may quickly improve access, it does carry some risk of local monopolies.

New technology, including the Internet, smart cards, and the use of mobile phones, can help to broaden access, although it does not necessarily address the underlying distortions limiting access (see BIS 2004, for a general overview of e-finance developments). On one end of the income spectrum, in Vienna payments for parking fees and in Finland payments at vending machines can be made by mobile phone. In many industrial countries electronic payments can be made through voice access, text messaging (SMS), or wireless application protocol (as a gateway to the Internet). Another arrangement in industrial countries allows customers to pay using the prepaid value stored on their mobile phone as a direct debit or to pay later, with charges for goods or services placed on the customer's phone bill. Use of mobile phones for financial services provision might facilitate access in developing countries, where mobile phones are often more widespread than fixed lines and can have a lower threshold for many users than banks do.

In some developing and transition countries (Bolivia, Brazil, China, Ghana, India, Lithuania, Malawi, Malaysia, Mexico, Nigeria, the Philippines, Russia, Turkey, and Venezuela; BIS 2004), banks have offered prepaid cards that can facilitate payment services for low-income households. Often, though, this requires regulatory changes. Technology can help in other ways. Hand-held remote transaction tools are being used by several microfinance institutions to process on the spot loan applications and approvals. In Uganda Hewlett-Packard and other technology firms active in the microfinance industry have been working to increase the scale of microfinance. They have developed a remote transaction system using hand-held devices to capture transaction data and transmit it back to management information systems on head office servers. (See [www.microsave.org](http://www.microsave.org) for other examples.)

In Mexico, in a program developed by Nafin, a government development bank, many small suppliers use their receivables from large creditworthy buyers, including foreign multinationals, to receive working capital financing. By effectively transferring the creditworthiness of large firms to small firms, the program allows small firms to access more and cheaper financing. Nafin operates an Internet-based platform, reducing costs, increasing transparency, and improving security. In the short run overhead costs are being subsidized, but by lowering costs for working capital for small firms, the program expects to generate more business and become sustainable (Klapper 2004).

Standard Bank of South Africa has also tried new ways of meeting the needs of an unbanked population (Paulson and McAndrews 1998). In 1993 Standard Bank set up E-Bank, offering card-only access to a simple savings product. It was supported by a dedicated staff speaking a mix of local languages and operating out of dedicated outlets to help overcome problems of illiteracy and concerns about security in a high crime environment. It had high start-up costs but provided financial services to a

low-income segment of the population. E-bank has since been absorbed in the bank's more general provision of financial services to low-income households.

Many other examples could be cited of specific market approaches and government interventions to enhance access. More generally, there has been much emphasis recently on facilitating the mainstreaming of microfinance institutions using traditional banking approaches and scaling up new initiatives on access. These initiatives can be implemented through specific interventions, as the above examples and work under way in India (Ananth and Mor 2005; Basu and Srivastava 2005) and other places show, but how to generalize is still a lesson to be drawn.

## Government Interventions to Broaden Access

The discussion so far has shown that it is not easy to determine how much of the failure to achieve universal use of financial services is due to lack of demand and how much to lack of supply. As with other goods and services, so for financial services demand may not exist even when access does. Many households choose not to have a bank account as they write no checks, collect their wages in cash, and transact their finances in cash. So, while they likely have access, they may not be burdened by lack of use. Firms that do not use external credit may choose not to do so because their rates of return on capital are too low to justify formal finance or because they are not willing to provide the necessary information about their business to banks and by implication to others, including the tax authorities. Equally important, and even in the best financial systems, financial service providers may not wish to supply financial services to all customers because it is not profitable or sustainable to do so. This does not reflect any market failures, but rather that finance, like other services, has its own demand and supply forces. This may simply mean that a country requires a certain overall level of development before more universal use is a viable proposition.

More generally, the poor and disenfranchised do not use financial services may be more a problem of poverty than of access. Although data are weak and do not allow definitive assessment, there are likely many people among the group with no access in developing countries who have no demand for financial services. Consequently, the share of those with potential demand for financial services but no access in developing economies may well be small and similar in size to that in industrial countries. Because there is evidence that use rises with per capita income and wealth, although with complex causality links, arguably the focus should primarily be on poverty-reducing growth and programs to enhance overall inclusion (jobs, education, and social participation), with greater use of financial services to follow as a corollary.

To determine whether there is a case for universal provision of financial services, more needs to be known about the benefits of access, about why households and

firms demand (or do not) financial services, about why financial service providers provide (or do not) financial services, and about the costs to society of providing greater access.

Much remains unknown. We do not know at the microlevel sufficiently well what the benefits and impacts of finance are. The gains of access to basic health care services such as immunization are much easier to document than the gains from access to financial services. There is also evidence that, from a social point of view, people invest too little in primary health care or education, thus justifying government intervention. We do not know systemically, however, whether individuals underuse basic (formal) financial services even when they have access at a reasonable cost.<sup>10</sup> Furthermore, access to credit may be a problem when it leads to impoverishing indebtedness from overborrowing. There is plenty of anecdotal evidence that some households may have difficulty managing access to credit, suggesting that some restraint in the use of financial services, say until financial literacy is more adequate, may well be welfare enhancing.

More generally, little is known about whether there is a public goods argument to be made in favor of extending access more broadly. There is a general poverty trap literature that highlights the key role of critical thresholds of consumption and investment in perpetuating poverty in the absence of functioning financial markets (see Azariadis 2005 for a review). This strand of analytical work, however, has not yet focused on the issue of desirable government interventions in financial services provision. There is work on other services, notably telecommunications and postal services, which have some closer parallels. As with these services, financial services display some properties of network industries (Claessens and others 2003). There are fixed costs on investment in branches and externalities of use as in payment systems and stock markets, where additional use lowers the unit costs. Payment systems, branches, and automated teller machines and other points of sale are distribution networks, similar to telecommunication networks and post offices, and may have parallels to these industries in arguments for and against government intervention. Also, as these industries typically have universal service objectives, there may be lessons on the preferred ways in which government can intervene to broaden access (for example, by subsidizing the user or the provider or through universal service obligations).<sup>11</sup>

Trying to broaden access, as will be clear by now, should not necessarily be a public policy goal. Public interventions, if any, will need to be carefully considered. Given political economy factors, broadening access may not relax credit and savings constraints when there is selection bias—when households or firms with good prospects and possibly already having access apply for credit. Subsidies not only distort markets, but evidence is mounting that subsidies are captured by the relatively well off, who often already have access. Priority lending requirements can also divert resources from the lowest-income segments. For example, interventions to improve

the supply of housing finance often end up being a subsidy for the middle class. In Brazil the cost of the housing finance program is an important factor behind the generally high financial intermediation spreads, hurting borrowers and depositors through higher lending rates and lower deposit rates, especially those less well off. In the end enhancing access can hurt those truly in need.

Another example of possibly perverse interventions relates to microfinance institutions. Multilateral financial institutions and bilateral donors have given much emphasis to microfinance institutions, including providing subsidies for setting up such institutions. These subsidies can work perversely, leading to higher subsequent spreads to recover fixed costs (Hoff and Stiglitz 1998). Thus, direct and indirect subsidies should remain minimal, and cost and risk cosharing with the private sector are important market tests. And even where there is a case to try to extend financial services provision to a larger segment of the population, it may be that the costs of such provision outweigh its benefits, as when the means to raise the necessary fiscal revenues are very distortive.

There is some evidence that the demand for and supply of financial services may be stimulated in other, less costly ways. Many employers prefer to deposit their payroll and wages electronically and would be willing to provide some form of subsidy to encourage use of formal bank services by employees (for example, facilitating branching within the premises, encouraging the establishment of a credit union, or facilitating private savings schemes). Governments can also do this. They can, for example, try to expand electronic transfers of social security, tax, and other individual-oriented payments to encourage more bank access.

For example, in 1999 the U.S. Treasury Department initiated a program to pay all federal benefits, such as social security payments, by electronic transfer accounts. One impediment was the large number of recipients without bank accounts who cashed their checks instead of depositing them in a bank account. Subsidies were used to encourage banks to open accounts and recipients to switch to electronic payments. The Treasury offered to pay banks \$12.60 for each electronic transfer account established for benefit recipients and specified a minimum set of characteristics that these accounts must meet (the accounts could not cost account owners more than \$3 a month and banks could not levy a fee for electronic deposits coming in). The switch would benefit the government as supplier (lower costs) and could also help recipients by inducing them to use financial services. In the end the take-up was less than expected, suggesting again that lack of use reflects lack of demand rather than lack of access and is part of a broader issue of social exclusion.

There are other options open for governments to stimulate use of banking and other basic financial services by households. For one, the regulatory system can be used to direct, but not mandate, banks to address the problem. This might be described as the Community Reinvestment Act (CRA) approach, following the model used in the United States. Second, authorities can mandate all banks to provide

minimum banking services (basic accounts) for otherwise excluded segments of the market. Third, governments can rely on banks with a social commitment (in the legal form of public banks, cooperatives, foundations, the postal network, or proximity banks such as local savings banks) to offer very restricted retail services. Each approach has advantages and disadvantages.<sup>12</sup>

The U.S. CRA, enacted in **1977** and revised in **1995**, aims to help meet the credit requirements of the communities in which banks operate, including low- to moderate-income neighborhoods. Each bank is rated every three years on its performance in making loans to low- and moderate-income people, allowing the public to apply pressure for noncompliance. Ratings focus on lending, services, and investment, with lending carrying the most weight. Claims for its success are contested, with neither side establishing a strong position. The CRA model is very specific and has not been followed elsewhere, which suggests that its replicability is limited. The CRA should not be seen in isolation but within the broader, political economy context of exclusion.

France, Ireland, Sweden, and the United Kingdom, among others, have tried to broaden access by legal means. In France anyone who applies to open a bank account but is rejected can contact the Bank of France, which will provide a bank (often the postal bank) that will be obligated to open an account. In some countries postal banks (often government owned) are required to provide basic cash and banking services. There is little review of the effects and efficiency of these schemes, however. Peachey and Roe (2004) review experience with proximity banks and find some support for a positive effect on access from a greater presence of such banks. Also, credit unions and other not for profit financial institutions can make a difference in access.

Credit extension programs, especially for small and medium-size enterprises, have been plentiful in both industrial and developing countries, suggesting a large public need to provide these forms. The efficacy of these interventions is doubtful, however, and the need seems to have arisen largely from political economy pressures (for a general review of credit lines, see Caprio and Demirgüç-Kunt **1997**; World Bank 2005c that provide some empirical evidence on subsidies and review general experiences). The means of distributing credit under these programs is generally distortive, credit often does not reach the intended target group but rather the well-connected, and institutional development is undermined, as banks do not develop their credit analysis skills. The case for direct and indirect intervention in access to credit is therefore less clear than the case for access to basic savings, payments, and transaction services.

## Conclusions

Over the last decade finance has been recognized as an important driver of economic growth. More recently, access to financial services has been recognized as an important

aspect of development, and more emphasis is being given to extending financial services to low-income households. Although analysis is just beginning, there is some evidence that access is improving. On the household side there are some data on the use of microfinance that suggest an expansion of use by households. Data have to be interpreted carefully, however, as increases may represent better data coverage over time rather than expansion. There is also evidence of more mainstreaming of financial service provision by commercial banks as competitive forces and technology allow them to reach lower-income segments of the population. Examples in developing countries are ICICI bank and the SHG Bank Linkage program in India and commercial banks in South Africa that have made it a priority to reach out to lower-income groups.

For firms the evidence on access to credit is more mixed. It appears to be increasing in some countries, but mostly in consumer finance forms, and less so on the small- and medium-size enterprise credit side. Some analysts have argued that recent trends in banking systems may have adverse consequences. Consolidation of the banking system in many countries increases the distance between borrower and lender, so that lending is based more on hard information, reducing the role of relationships, which can be especially useful for new and small firms. Yet part of this increased consolidation is a consequence of increased competition, which in general helps to increase access. Indeed, while there is cross-country evidence that more concentrated banking systems could increase financing obstacles, this is more so if the system is not competitive and is dominated by public banks.

A more definite interpretation of the factors affecting access will have to await better data on access and use at both the micro and the macro level. This will require actions by national and international agencies to develop more comparable data on use and access barriers. At the micro level there has already been more emphasis in recent years on monitoring and evaluation using impact data, including on access to financial services, by donors, the Consultative Group to Assist the Poor, the International Finance Corporation, and others. These data are often not comparable, however. Furthermore, as policy moves away from specific lending and other interventions and emphasizes the general policy and institutional environment, there is more need to measure access to financial services at the system level as well.

Data on use will have to come from different sources: providers of financial services (directly and from national statistics), users of financial services (from surveys), and experts (to identify constraints). Each of these data sources has tradeoffs—in quality, costs, and coverage—so simultaneous actions will be needed. Without better data, however, little progress can be made on policy recommendations (Honohan 2005; World Bank 2005b). In addition, data can be collected across countries on the terms and conditions under which financial services are being provided—costs, type of services, requirements—to provide some insights into barriers to access. With better data and benchmarking systems

(across and over time), more analysis on what is driving use and better identification of the barriers to access will be possible.

In addition, analysis of the success of different models aimed at enhancing access and rigorous empirical evaluations of government interventions are needed. Controlled "experiments," such as those by Karlan and Zinman (2005), in which consumers were randomly offered different terms on possible loans, can provide good insights into the functioning of credit markets. By applying different treatments to different forms of financial service provision (for example, by introducing new technology "randomly" at the branch level), it may be possible to better distinguish which reforms aimed at enhancing access are most successful in what circumstances (the Centre for Micro Finance Research in India, the Poverty Action Lab at MIT, and others are pioneering research in this area). This type of analysis will help private financial institutions deliver financial services profitably and guide national and international policy interventions. It might also be useful for international and national agencies to continue to develop "models" on various new aspects of access, such as advice on regulations of microfinance institutions and their activities as the new Basle Accord is being implemented, and rules for some aspects such as consumer protection, know your customer rules, and anti-money laundering and counterterrorism financing. And, finally, guidance on what data to collect, how and from whom will be necessary.<sup>13</sup>

## Notes

Stijn Claessens is a senior adviser in the Operations and Policy Department, Financial Sector Vice Presidency, at the World Bank. The author can be reached via e-mail at [sclaessens@worldbank.org](mailto:sclaessens@worldbank.org). This article is based on a paper presented at the Fifth Services Experts Meeting on Universal Access, Paris, February 3-4, 2005, organized by the Organisation for Economic Co-operation and Development and the World Bank and at the Centre for Economic Policy Research Economic Policy Roundtable, London, July 15, 2005. It draws on material by Stephen Peachey and Alan Roe and presentations by Priya Basu, Thorsten Beck, Asli Demirgüç-Kunt, Patrick Honohan, Anjali Kumar, Tova Solo, and others at the World Bank working on access to financial services. The author thanks Thorsten Beck, Antonio Estache, Asli Demirgüç-Kunt, Aaditya Mattoo, meeting participants, the two referees, and the editor for helpful comments and Ying Lin for research assistance.

1. Although many empirical studies of growth find a positive role for financial development, not all do. Some questions remain on causality and on missing or omitted variables. For example, Bosworth and Collins (2003) do not find a statistically significant relationship between financial sector development and growth. Of course, these and other regression results depend on what is included as explanatory variables, and although extensive robustness tests have been conducted, definite answers may remain elusive.

2. See IMF and World Bank (2005) on the Millennium Development Goals and progress in achieving them.

3. Beck, Demirgüç-Kunt, and Maksimovic (2004) investigate the use of different financing sources across countries and find that financial and institutional development are associated with higher of formal external financing sources, especially for small firms. See also Beck and others 2005 and Ayyagari, Demirgüç-Kunt, and Maksimovic 2005.

4. This group of those with no access could be further split into (a) those who want to use financial services and have no access and (b) those who do not want to use financial services and have no access.
5. This requires answering the question of what constitutes "reasonable cost," where reasonable has to be defined relative to some objective standard, and costs have to reflect all forms of pecuniary and nonpecuniary costs. In an application to South African countries a cutoff level of a maximum cost of 2 percent of income was used to deem financial services to be accessible (Genesis Analytics 2004; see also Napier 2005).
6. For example, Kempson and others 2000 distinguish five types of exclusion from financial services.
7. These two works used the World Bank Economic Survey data; the World Bank Investment Climate Assessment data are currently being analyzed to further determine what drives lack of access to external financing for firms across countries.
8. See some of the papers presented at a recent World Bank conference on small- and medium-size enterprises, [www.worldbank.org/research/projects/sme\\_conference.htm](http://www.worldbank.org/research/projects/sme_conference.htm).
9. This is not to say that there should be a free flow of information. Concerns about privacy warrant some restrictions in the sharing of financial information about households across different types of institutions and even within a single financial institution. Furthermore, if institutions do not obtain some return from their information, their incentives to collect reliable information will be limited.
10. There is evidence, however, that people do not save sufficiently for their old age, but that refers more to a quantity dimension. And there is the evidence that some people prefer to use informal financial services even when these are more costly because of other, nonprice barriers for formal services.
11. Work by Jean Tirole, Jacques Cremers, and others at the University of Toulouse has started to address the issue of regulation of networks such as in payments system and credit cards; see Claessens and others (2003) for a review.
12. The following sections draw extensively on Peachey and Roe, 2004.
13. See World Bank (2005a) for a first attempt to develop core indicators in household surveys.

## References

- Ananth. Bindu, and Nachket Mor. 2005. "Financial Services Case Study: India." Paper presented at the Organisation for Economic Co-operation and Development and World Bank's Fifth Services Experts Meeting on Universal Access, February 3-4, Paris.
- Ayyagari, Meghana, Asli Demirgüç-Kunt, and Vojislav Maksimovic. 2005. "How Important Are Financing Constraints? The Role of Finance in the Business Environment." Policy Research Working Paper 3820. World Bank, Washington, D.C.
- Azariadis, Costas. 2005. "Poverty Traps." In Steven Durlauf and Phillipe Aghion, eds., *Handbook of Economic Growth*. Amsterdam: North-Holland Elsevier Publishers.
- Barr, Michael. 2005. Microfinance and Financial Development. *Michigan Journal of International Law* 26(1): 271-96.
- Barth, James, Gerard Caprio, Jr., and Ross Levine. 2005. *Rethinking Bank Regulation and Supervision: Till Angels Govern*. New York: World Bank and Cambridge University Press.
- Basu, Priya, and Pradeep Srivastava. 2005. "Scaling-up Access to Finance for India's Rural Poor." Policy Research Working Paper 3646. World Bank, South Asia Region, Finance and Private Sector Development Unit, Washington, D.C.
- Beck, Thorsten, and Augusto de la Torre. 2005. "Broadening Access to Financial Services: Risks and Costs." PowerPoint Presentation. World Bank, Washington, D.C.
- Beck, Thorsten, Asli Demirgüç-Kunt, and Maria Soledad Martinez Peria. 2005. "Reaching Out: Access To and Use of Banking Services Across Countries." Policy Research Working Paper 3754. World Bank, Washington, D.C.

- Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine. 2004. "Finance, Inequality and Poverty: Cross-Country Evidence." Policy Research Working Paper 3338. World Bank, Washington, D.C.
- . 2005. "SMEs, Growth, and Poverty: Cross-Country Evidence." *Journal of Economic Growth* 10(3): 199–229.
- Beck, Thorsten, Asli Demirgüç-Kunt, and Vojislav Maksimovic. 2004. "Financing Around the World: Are Small Firms Different?" Policy Research Working Paper 2905. World Bank, Washington, D.C.
- . 2005. "Financial and Legal Constraints to Firm Growth: Does Firm Size Matter?" *Journal of Finance* 60(1): 137–77.
- Beck, Thorsten, Ross Levine, and Norman Loayza. 2000. "Finance and the Sources of Growth." *Journal of Financial Economics* 58(1–2): 261–300.
- Beck, Thorsten, Asli Demirgüç-Kunt, Luc Laeven, and Ross Levine. 2005. "Finance, Firm Size, and Growth." Policy Research Working Paper 3485. World Bank, Washington, D.C.
- Beck, Thorsten, Asli Demirgüç-Kunt, Luc Laeven, and Vojislav Maksimovic. 2004. "The Determinants of Financing Obstacles." Policy Research Working Paper 3204. World Bank, Washington, D.C.
- Berger, Allen, and David Humphrey. 1997. "Efficiency of Financial Institutions: International Survey and Directions for Future Research." *European Journal of Operational Research* 98(2): 175–212.
- Berger, Allen, and Greg Udell. 2005. "A More Complete Conceptual Framework for SME Financing." Policy Research Working Paper 3795. World Bank, Washington, D.C.
- BIS (Bank for International Settlements). 2004. "Survey of Developments in Electronic Money and Internet and Mobile Payments: Information on Current and New Innovative Products." Consultative Report. Bank for International Settlements, Basle.
- Board of Governors, U.S. Federal Reserve System. 2004. *Survey of Consumer Confidence 2004*. Federal Reserve Board, Washington, D.C.
- Bodie, Zvi, and Robert Merton. 1995. "A Conceptual Framework for Analyzing the Financial Environment." In Dwight B. Crane, Zvi Bodie, Kenneth A. Froot, Andre F. Perold, and Robert C. Merton, eds., *The Global Financial System: A Functional Perspective*. Boston, Mass.: Harvard Business School Press.
- Bosworth, Barry, and Susan Collins. 2003. "The Empirics of Growth: An Update." *Brookings Papers on Economic Activity* 2: 113–206.
- Caprio, Gerard, Jr., and Asli Demirgüç-Kunt. 1997. "The Role of Long Term Finance: Theory and Evidence." Policy Research Working Paper 1746. World Bank, Washington, D.C.
- Caskey, John, Clemente Ruíz Durán, and Tova Solo. 2004. "The Unbanked in Mexico and the United States." Paper presented at the International Conference on Access to Finance, World Savings Banks Institute in association with the World Bank, October 28–29, Brussels.
- CGAP (Consultative Group to Assist the Poor). 2004. "Financial Institutions with a 'Double Bottom Line': Implications for the Future of Microfinance." Occasional Paper 8. Consultative Group to Assist the Poor, Washington, D.C.
- Christen, Robert Peck, Timothy R. Lyman, and Richard Rosenberg. 2003. *Microfinance Consensus Guidelines: Guiding Principles on Regulation and Supervision of Microfinance Institutions*. Washington, D.C.: Consultative Group to Assist the Poor.
- Claessens, Stijn, and Enrico Perotti. 2005. "The Links between Finance and Inequality: Channels and Evidence." Background paper for *World Development Report 2006*. World Bank, Washington, D.C.
- Claessens, Stijn, Gergely Dobos, Daniela Klingebiel, and Luc Laeven. 2003. "The Growing Importance of Networks in Finance and Its Effects on Competition." In Anna Nagurney, ed., *Innovations in Financial and Economic Networks*. Northampton, Mass.: Edward Elgar Publishers.
- Clarke, George, Lixin Colin Xu, and Heng-fu Zou. 2003. "Finance and Income Inequality: Test of Alternative Theories." Policy Research Working Paper 2984. World Bank, Washington, D.C.

- Clarke, George, Robert Cull, and Maria Soledad Martinez Peria. 2001. "Does Foreign Bank Penetration Reduce Access to Credit in Developing Countries? Evidence from Asking Borrowers." Policy Research Working Paper 2716. World Bank, Washington, D.C.
- Clarke, George, Robert Cull, Maria Soledad Martinez Peria, and Susana M. Sanchez. 2003. "Foreign Bank Entry: Experience, Implications for Developing Economies, and Agenda for Further Research." *World Bank Research Observer* 18(1): 25–59.
- De la Torre, Augusto, Juan Carlos Gozzi, and Sergio Schmukler. 2005. "Innovative Experiences in Access to Finance: Market Friendly Roles for the Visible Hand?" *Latin America Regional Studies Series*. World Bank, Washington, D.C.
- Demirgüç-Kunt, Ash, and Ross Levine, eds. 2001. *Financial Structure and Economic Growth: A Cross-Country Comparison of Banks, Markets, and Development*. Cambridge, Mass.: MIT Press.
- Demirgüç-Kunt, Ash, and Vojislav Maksimovic. 1998. "Law, Finance, and Firm Growth." *Journal of Finance* 53(6): 2107–37.
- Emerging Market Economics. 2005. "Financial Access Indicators Stocktake." Prepared for the Department for International Development. London.
- Francisco, Manuela, and Anjali Kumar. 2005. *Enterprise Size, Financing Patterns, and Credit Constraints in Brazil: Analysis of Data from the Investment Climate Assessment Survey*. Washington, D.C.: World Bank.
- Gasparini, Leonardo, Alejandro Tamola, Federico Gutiérrez, Leopoldo Tornarolli, and Guido Porto. 2005. "Finance and Credit Variables in Household Surveys of Developing Countries." World Bank, Washington, D.C.
- Genesis Analytics. 2004. "A Survey of the SADC Region: South African Financial Institutions, Regional Policies and Issues of Access." Genesis Analytics, Johannesburg, South Africa.
- Glaessner, Thomas, Nagavelli Anamali, Stijn Claessens, Karen Furst, Daniela Klingebiel, Leora Klapper, and Robert Schware. 2004. "South Africa: Technology and Access to Financial Services: Lessons from Experience." World Bank, Washington, D.C.
- Halac, Marina, and Sergio Schmukler. 2003. "Distributional Effects of Crises: The Role of Financial Transfers." Policy Research Working Paper 3173. World Bank, Washington, D.C.
- Hanson, James, Patrick Honohan, and Giovanni Majnoni. 2003. *Globalization and National Financial Systems*. Washington, D.C.: World Bank and Oxford University Press.
- Hoff, Karla, and Joseph E. Stiglitz. 1998. "Moneylenders and Bankers: Price-increasing Subsidies in a Monopolistically Competitive Market." *Journal of Development Economics* 55(2): 485–518.
- Honohan, Patrick. 2004. "Financial Sector Policy and the Poor: Selected Findings and Issues." World Bank Working Paper 43. Washington, D.C.
- . 2005. "Measuring Microfinance Access: Building on Existing Cross-Country Data." Policy Research Working Paper 3606. World Bank, Washington, D.C.
- IMF (International Monetary Fund) and World Bank. 2005. *Global Monitoring Report: From Consensus to Momentum*. Washington, D.C.
- Jansson, Tor, Ramon Rosales, and Glenn D. Westley. 2004. *Principles and Practices for Regulating and Supervising Microfinance*. Washington, D.C.: Inter-American Development Bank.
- Karlan, Dean S., and Zinman, Jonathan. 2005. "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment." Working Paper 911. Economic Growth Center, Yale University, New Haven.
- Kempson, Elaine, Claire Whyley, John Caskey, and Sharon Collard. 2000. "In or out? Financial Exclusion: A Literature and Research Review." Consumer Research Report 3. Financial Services Authority, London.
- Klapper, Leora. 2004. "The Role of Reverse Factoring in Supplier Financing of Small and Medium Sized Enterprises." Policy Research Working Paper 3593. World Bank, Washington, D.C.

- Klapper, Leora, Luc Laeven, and Raghuram Rajan. 2004. "Business Environment and Firm Entry: Evidence from International Data." Policy Research Working Paper 3232. World Bank, Washington, DC.
- Kumar, Anjali. 2005. *Access to Financial Services in Brazil: A Study*. Directions in Development. Washington, DC.: World Bank.
- Kumar, Anjali, Thorsten Beck, Cristine Campos, and Soumya Chattopadhyay. 2004. "Assessing Financial Access in Brazil." World Bank Working Paper 50. Washington, DC.
- Levine, Ross. 2005. "Finance and Growth: Theory, Evidence, and Mechanisms." In P. Aghion and S. Durlauf, eds., *Handbook of Economic Growth*. Amsterdam: North-Holland Elsevier Publishers.
- Littlefield, Elizabeth, Jonathan Morduch, and Syed Hashemi. 2003. "Is Microfinance an Effective Strategy to Reach the Millennium Development Goals?" CGAP FOCUS Note 24. Consultative Group to Assist the Poor. Washington, DC.
- Maimbo, Samuel Munzele, and Dilip Ratha, eds. 2005. *Remittances, Development Impact and Future Prospects*. Washington, DC.: World Bank.
- Miller, Margaret. 2001. *Credit Reporting Systems and the International Economy*. Cambridge, Mass.: MIT Press.
- Morduch, Jonathan. 1999. "The Microfinance Promise." *Journal of Economic Literature* 37(4): 1569–1614.
- . Forthcoming. "Microinsurance: The Next Revolution?" In Abhijit Banerjee, Roland Benabou, and Dilip Mookherjee, eds., *What Have We Learned About Poverty?* New York: Oxford University Press.
- Morduch, Jonathan, and Barbara Hayley. 2002. "Analysis of the Effects of Microfinance on Poverty Reduction." NYU Wagner Working Paper 1014. New York University, New York.
- Napier, Mark. 2005. "Provision of Financial Services in South Africa." Paper presented at the Organisation for Economic Co-operation and Development and World Bank's Fifth Services Experts Meeting on Universal Access, February 3–4, Paris.
- Orozco, Manuel. 2004. *Remittances to Latin America and the Caribbean: Issues and Perspectives on Development*. Washington, DC.: Organization of American States.
- Paulson, Joanne, and Jaime McAndrews. 1998. "Financial Services for the Urban Poor—South Africa's EPlan." Working Paper 449. World Bank, Washington, DC.
- Peachey, Stephen, and Alan Roe. 2004. "Access to Finance." A Study for the World Savings Banks Institute, Oxford.
- Pesaresi, Nicola and Odile Pilley. 2003. "Retail Banking, Social Exclusion, and Public Service." *EU Competition Policy Newsletter* 1: 14–20.
- Rajan, Raghuram, and Luigi Zingales. 1998. "Financial Development and Growth." *The American Economic Review* 88(3): 559–86.
- . 2003. *Saving Capitalism from the Capitalists: Unleashing the Power of Financial Markets to Create Wealth and Spread Opportunity*. New York: Crown Business.
- Sanchez, Susana M., George Clarke, Robert Cull, and Maria Soledad Martinez Peria. 2005. "Bank Lending to Small Businesses in Latin America: Does Bank Origin Matter?" *Journal of Money Credit and Banking* 37(1): 83–118.
- Solo, T. M. 2005. "Broadening Access to Financial Services Among the Urban Population: Mexico City's Unbanked." Sector Report 32418. Washington, DC.: World Bank.
- Stiglitz, J. E. and A. Weiss. 1981. "Credit Rationing in Markets with Imperfect Information." *American Economic Review* 71(3): 393–410.
- World Bank. 2001. *Finance for Growth, Policy Choices in a Volatile World*. New York: Oxford University Press.

- . 2005a. "Core Concepts and Definitions for Measuring Financial Access in Households Surveys." Report of a Working Group on Harmonizing Households Surveys on Access to Finance. Office of the Vice President, Financial Sector, Washington, D.C.
- . 2005b. "Measuring Financial Access, Outlining the Scope of Current Data Collection Efforts." Report of a Working Group. Office of the Vice President, Financial Sector, Washington, D.C.
- . 2005c. "Review of Bank Lending for Lines of Credit." Report of a Working Group. Operations Evaluation Department, Washington, D.C.
- . 2006. "Provision of Financial Services to the Rural Poor: What Can Be Done When Services Aren't Available?" Version 1.6. In *Scaling Up Poverty Reduction: Case Studies and Global Learning Process*. Proceedings of May 25–27, 2004, Shanghai Conference, Washington, D.C.

---

# What Have We Learned from a Decade of Manufacturing Enterprise Surveys in Africa?

---

Arne Bigsten • Måns Soderbom

---

*In the early 1990s the World Bank launched the Regional Program on Enterprise Development (RPED) in several African countries, a key component of which was to collect data on manufacturing firms. The data sets built by these and subsequent enterprise surveys in Africa generated considerable research. This article surveys the research on the African business environment, focusing on risk, access to credit, labor, and infrastructure, and on how firms organize themselves and do business. It reviews the research on enterprise performance, including enterprise growth, investment, and exports. The article concludes with a discussion of policy lessons.*

The idea that better private-sector performance in Africa can reduce poverty remains central in policy discussions. Even though the performance of Africa's manufacturing sector has generally been quite poor over the last 10–15 years (Mauritius is a conspicuous exception), many people still believe that manufacturing can act as an engine of growth on the continent, by creating skilled jobs and positive spillover effects and, more generally, by modernizing the economy. In the early 1990s, to get a better idea of why things were going wrong in the manufacturing sector and how to improve them, the World Bank fielded extensive data collection projects in many African countries. Such data collection has subsequently become a regular exercise in at least some African countries. This article reviews the findings that have emerged from this research, focusing primarily on research published after peer review.'

The results reported in this survey are thus based on enterprise survey data. The article argues that the availability of such data has greatly improved our understanding of Africa's manufacturing sector. Certain research questions cannot be answered by relying on aggregate data. It may be that the researcher has to analyze

differences across firms within an economy to estimate, say, productivity dispersion—a conventional indicator of the degree of competition—or to find out whether credit-constrained firms invest less than unconstrained firms. It is arguably in cases like this, when heterogeneity across firms is of central interest, that firm data are most useful. The more heterogeneity, the more pertinent are questions relating to the causes and consequences of such differences across firms, and the more pressing, the need for enterprise data. Panel data, which follow the same firms over time, are especially powerful in this context. They permit drawing stronger conclusions about causal effects, because time-invariant unobserved heterogeneity can be controlled for. As shown below, panel data also provide a good basis for analyzing firm dynamics, for example, productivity growth and firm survival. One of the recurring themes of this article is that enterprise data indicate substantial heterogeneity in choices and outcomes across firms within countries in Africa.<sup>2</sup> Indeed, while most African firms have not fared well during the last decade, some have performed extremely well.

In recent years there has been a rapidly growing interest in documenting differences in the investment climate across regions and countries and investigating whether such differences can be linked to differences in outcomes across and within countries (Batra, Kaufmann, and Stone 2003; World Bank 2004, 2005).<sup>3</sup> Survey data have played a key role in this context. It is often argued that the poor investment climate in Africa results in high transaction costs and particularly disadvantages the manufacturing sector and its ability to export, because manufacturers are intensive users of investment climate services (Collier 2000). This is problematic because exporting may present a promising route to growth and development in Africa. Domestic markets for manufactured goods are typically very small, so if African industry is to grow, generate new jobs, and contribute to poverty reduction, a substantial share of its output needs to be oriented toward exports. There is also evidence from the survey data that firms become more productive as a result of exporting. This is the so-called learning by exporting effect, first documented for Africa by Bigsten and others (2004) and subsequently confirmed on a larger sample by Van Biesebroeck (2005a). This is discussed in more detail below.

The following section briefly describes the main data collection projects that have been fielded in Africa since the early 1990s and comments on the quality of the data. The article then reviews the research results on manufacturing firms in Africa, looking first at the business environment, especially at risk, access to credit, labor and skills, and infrastructure. It turns next to how firms that operate in such an investment climate organize themselves and how firms do business. It looks in some detail at what has been learned about three key aspects of firm performance: growth, investment and technology acquisition, and exports. The final section presents some lessons.

## African Enterprise Survey Data

In the last 10–15 years five major research programs have collected survey data on African enterprises on a large-scale. The first was the Regional Program on Enterprise Development (RPED), which carried out surveys in Burundi, Cameroon, Côte d'Ivoire, Ghana, Kenya, Tanzania, Zambia, and Zimbabwe between 1992 and 1995. Each survey typically covered about 200 firms, and the same firms were surveyed three years in a row in most countries, thus yielding panel data. The surveys covered large and small firms, including informal firms, in the food, wood, textiles, and metal industries. Data were collected on a wide range of variables, including sales and output, capital stock, entrepreneur characteristics, employment by occupational category, labor turnover, wages, and management–worker conflict.

Data collection as part of the RPED slowed for a while after 1995. The World Bank subsequently initiated the Firm Analysis and Competitiveness Surveys (PACS), a large data collection program oriented toward larger firms outside sub-Saharan Africa. Morocco was the only African country surveyed. However, at the end of the 1990s and in the early 2000s, new survey initiatives took place in sub-Saharan Africa, as Investment Climate Surveys and the World Business Environment Survey (WBES) were fielded across a wide range of countries. The Investment Climate Surveys carried out so far in sub-Saharan Africa have covered Cameroon, Côte d'Ivoire, Eritrea, Ethiopia, Kenya, Mauritius, Mozambique, Nigeria, Senegal, Tanzania, Zambia, and Zimbabwe. These surveys were similar to the earlier RPED surveys in firm and sector coverage, but the survey instruments were oriented more toward investment climate issues, such as governance, regulations, economic policy, and public services. The WBES was launched in 80 countries and one territory in 1999 and 2000, including 17 countries in sub-Saharan Africa (Batra, Kaufmann, and Stone 2003). The surveys in sub-Saharan Africa covered 52–137 firms per country and included both manufacturing and nonmanufacturing firms.

Another research program involving enterprise surveys was implemented by the Centre for the Study of African Economies (CSAE) at Oxford University. These surveys followed up on the RPED surveys in Ghana, Kenya, and Tanzania and have extended the panel data.<sup>4</sup> The panel data from the early RPED surveys include at most three observations per firm. The CSAE surveys have made some of these panels much longer. In the case of Ghana, for instance, the time-series dimension of the panel is currently 12 years.

The data sets generated by these surveys constitute the most comprehensive source of information on African manufacturing firms available. Still, several limitations need to be borne in mind when interpreting the results of the research based on these data. First, because large firms are typically oversampled, the samples generated by these surveys are generally not representative of the population of firms. And because sampling weights are typically either not available or are based on outdated

registers of firms (Van Biesebroeck 2005b), results based on the survey data may not generalize beyond the samples. Second, most results pertain to four industrial sectors (food, wood, textiles, and metal), and it is not clear whether the findings would apply to other sectors. Third, the quality of the data is not perfect. The financial data are likely to be particularly noisy. Consider the capital stock data, for example. Firms report the replacement value of plant and machinery, and sometimes they are asked to do this every year. Information on investment is also collected, so with two data points on capital and one on investment, the capital stock in period two should be approximately equal to the capital stock in period one (perhaps adjusted for depreciation) plus investment. In practice, however, there are sometimes large differences between the reported value of the capital stock and the value implied by last year's value plus investment. Measurement errors in the data may lead to serious biases.

## What Have We Learned?

This section looks at the business environment in which African manufacturing firms operate and then examines three key aspects of firm performance: growth, investment, and exports.

### *The Business Environment*

The business environment has emerged as the prime suspect for poor enterprise performance in Africa. Improving the investment climate is seen as a policy priority for the continent (World Bank 2004, 2005).<sup>5</sup> Batra, Kaufmann, and Stone (2003) analyze the obstacles to business based on the WBES data. The leading constraint cited by enterprise managers in Africa is financing, followed by corruption, infrastructure, and inflation. Pooling data across all regions, the researchers find a negative and statistically significant relationship between taxes, regulations, and financing, and the growth in sales and investment. Quantitatively, the largest effect is that of financing constraints on sales growth. Over a three-year period average sales growth is 4.6 percentage points lower for a firm identifying itself as financially constrained than for an unconstrained firm, everything else equal.

What can be inferred from these results? First, given the nature of the explanatory variables, it is important to distinguish between correlation and causality. There is almost certainly some bias caused by omitted variables or reverse causation. Financing, in particular, seems very likely to be endogenous: firms that perform badly and grow slowly will have difficulty getting financing. Batra, Kaufmann, and Stone (2003:71) acknowledge this and recommend interpreting the results as "empirical associations" rather than causal effects. Second, even if viewed as causal effects, the quantitative significance of the business environment effects is rather small compared

with the substantial variation in growth rates across firms in this data set. Third, the inclusion of country fixed effects makes the parameter estimates robust to omitted variables that vary across, but not within, countries. This could well be important. The fixed effects approach also implies the need for caution when comparing cross-country differences. The results are best interpreted as showing that a good *local* investment climate is good for *local* business performance. The results do not show whether firms in countries with poor average business conditions have lower or higher average growth rates of sales and investment than firms in countries with good average conditions.

One implication of a poor business environment is that the costs for certain services important to manufacturers will be high. Eifert, Gelb, and Ramachandran (2005) show that African firms have high indirect costs (transport, logistics, telecommunications, water, electricity, land and buildings, marketing, accounting, security, and bribes) compared with firms in Asia and that African firms suffer substantial losses from power outages, crime, shipment losses, and the like. Furthermore, economic risk in Africa is typically high, credit is expensive or unavailable, skilled labor is relatively expensive, and domestic markets are typically very small. It seems reasonable to suppose that these factors present significant problems for manufacturers in Africa. What do the enterprise data show about these characteristics of the business environment?

**Risk.** Entrepreneurs in Africa face considerable uncertainty with regard to prices (including foreign exchange), demand, customer payment, reliability of infrastructure, and corruption. Investigating the effects of risk empirically is difficult, however, because risk is not easily measured. In macroeconomic analysis, this is commonly done by using some measure of volatility in demand or the exchange rate. Gunning and Mengistae (2001) point out that this may be misleading, however, because the standard deviation of the exchange rate may not capture the credibility of the government's exchange rate policy. The African enterprise surveys, however, have included questions about expectations of macroeconomic variables (inflation, interest rates, and exchange rate) and enterprise variables (employment and output). Questions were asked, for example, about expectations of the extent of depreciation in the exchange rate during the next 12 months.

Such data have enabled researchers to get a much better understanding of the effects of perceived risk. Using data on Ghanaian manufacturing firms, Pattillo and Soderbom (2001) find that firms that face considerable uncertainty about future demand tend to have lower profit rates than firms facing less uncertainty. The researchers argue that high risk makes firms choose a conservative product mix, with a lower expected profit rate.

Fafchamps, Gunning, and Oostendorp (2000) show that Zimbabwean firms respond to risk by increasing their inventories, another example of how risk leads to

conservative behavior and additional costs. The most commonly cited effect of risk, however, is on investment. Gunning and Mengistae (2001) conclude from their review of the evidence that investments in African manufacturing have been held back by high risk rather than low returns on investments. This is discussed further below.

*Access to credit.* Bigsten and others (2003) look at formal credit market participation and credit constraints based on 1991–95 survey data on firms in Burundi, Cameroon, Côte d'Ivoire, Ghana, Kenya, and Zimbabwe. The researchers use data on firms' demand for external funds and on whether loan applications were approved. Summary data on the frequency of loan applications and outcomes suggest that demand for formal loans among African manufacturers is low: less than 20 percent of the firms in the sample had applied for a formal loan in the year before the survey (table 1). Among those applying, most firms obtained loans, but there are large differences by firm size. Loan applications are less common among small firms, and the success rate is lower than among large firms.

Of course, a firm may be credit constrained even if it does not apply for a loan. A firm may expect an application to fail precisely because there are credit constraints and may therefore decide to avoid the transaction costs and not apply. Based on information on why firms did not apply for a loan, Bigsten and others (2003) identify three groups of firms: those without credit demand and, among those with credit demand, constrained and unconstrained firms. Of the firms in their sample, 55 percent have no credit demand, 33 percent are credit constrained, and 12 percent are unconstrained (table 2). The differences are large across firms of different sizes. Close

**Table 1.** Formal Credit Market Participation by Firm Size (Percent)

	<i>Micro</i>	<i>Small</i>	<i>Medium</i>	<i>Large</i>	<i>All</i>
Did not apply	92	82	80	75	82
Applied and did not receive	6	11	9	5	8
Applied and received	2	7	11	20	10

*Source:* Bigsten and others (2003):table 2.

**Table 2.** Credit Constraints by Firm Size (Percent)

	<i>Micro</i>	<i>Small</i>	<i>Medium</i>	<i>Large</i>	<i>All</i>
No credit demand	33	50	67	66	55
Demand, but rejected <sup>a</sup>	64	42	21	10	33
Received loan	3	8	12	23	12

<sup>a</sup>Includes firms that suggested that a loan application would be rejected by banks.

*Source:* Bigsten and others (2003).

to two-thirds of the small firms appear constrained, but only 10 percent of the large firms. About two-thirds of the large firms choose not to participate in the credit market compared with only a third of the small firms. That the smallest firms are credit constrained is supported by regression results indicating that, controlling for other important factors such as expected profitability and indebtedness, the likelihood of a successful loan application varies with firm size. Although this suggests that banks are biased against small firms, the researchers note that this result may reflect banks' transaction costs. In any case, the size effect is substantial: for a small firm to have the same chance of getting a loan as a large firm, the small firm needs to have an average return on fixed capital more than 200 percentage points higher than the large firm.

In another study analyzing the links between formal credit and company investment, Habyarimana (2003) uses matched bank–firm data from Uganda to estimate the effect on firm performance of losing a banking relationship. Four Ugandan banks were closed between September 1998 and May 1999 because of imprudent banking practices. As a result, 30 percent of the firms in the sample lost one or several banking relationships. Habyarimana estimates that over the three years following the banking crisis, the average annual growth rate of employment among firms that lost a banking relationship was 2.3–4 percent lower than the average annual growth rate of unaffected firms, controlling for firm fixed effects and sector-specific time effects. Furthermore, firms affected by the banking crises are more likely to report being credit constrained, suggesting that losing a banking relationship hampers investment.

*Labor and skills.* Labor costs and the supply of labor in general and specific skills in particular are important for firm performance. Two general results in this area have emerged from the research on the African survey data, one related to earnings and education and one to earnings and firm size. The first is that earnings are positively correlated with education. Jones (2001) uses matched employer–employee data from Ghana to see whether wage differentials between workers with different levels of education reflect genuine productivity differentials. She estimates a productivity equation and an earnings equation and investigates whether the coefficients on education in the two equations are consistent with the hypothesis that earnings differentials reflect productivity differentials—whether the marginal product of labor is equal to the wage for given levels of education. In both equations the estimated coefficient on education is close to 0.07, and there is no statistically significant difference between the relative wages paid to workers and their relative productivity levels. This result, which appears robust to alternative specifications, supports the notion that education raises productivity.

Jones (2001) assumes that the return to education is constant. Soderbom and others (2006) relax this assumption and investigate whether the return to education

varies across different levels of education in Kenya and Tanzania during 1993–2001. It is often argued that the earnings function is concave in education, implying that the marginal return to education decreases with the level of education (Psacharopoulos 1994; Psacharopoulos and Patrinos 2002). The results reported by Soderbom and others (2006), however, indicate that the earnings function is convex, so that the marginal returns to education are lowest for individuals with the least education. This result is robust to endogeneity of education. One implication is that giving priority to investment in primary education may have little impact on poverty unless those affected proceed to higher levels of education. The researchers speculate that the convexity may be part of the explanation for why the rapid expansion of education in Africa has generated so little growth.

The second general result that has emerged from the research on enterprise data and labor is that wages differ significantly across different size firms, even when comparing workers with similar levels of human capital. In an early study documenting this, Velenchik (1997), using manufacturing data from Zimbabwe collected in 1993, estimates a wage-size elasticity of 0.26 for the sample of 1,609 observations. This indicates a very high size-wage gap. For instance, it implies that the average wage in a firm with 100 employees is about 80 percent higher than the average wage in a firm with 10 employees. The underlying regression does not control for other factors that might affect wages, so this result can be viewed as a reference point.

Velenchik (1997) goes on to investigate whether the size-wage gap can be attributed to omitted variables. She considers a very wide range of observable factors, including experience, tenure, education, training, gender, location, industry, occupation, various working condition variables, unionization, ownership, profitability, and productivity. Although the estimated wage-size elasticity falls as a result of controlling for these variables, it never drops below 0.10 and is always significantly different from zero. Thus, even conditional on this large set of observable variables, a quantitatively important size-wage gap remains.

It could be, of course, that the size-wage gap documented by Velenchik is driven by omitted *unobserved* factors. As she notes “[t]he basic labor quality difference argument is that larger firms hire ‘better’ workers, and therefore pay higher wages” (Velenchik 1997:311). If worker ability is partially unobserved, this would set up a correlation between size and wages. Recent evidence based on panel data suggests that this is not the case, however. Soderbom, Teal, and Wambugu (2005) show that, conditional on all observed and unobserved worker skills that are fixed over time, workers’ earnings in Ghana and Kenya vary positively with firm size. Again, the estimated effects are economically large. In the preferred specification for Ghana, the estimated elasticity of wages with respect to firm size is 0.15; in Kenya it is 0.08. The point estimates imply that as a firm grows from, say, having 20 employees to having 40, wages will increase 11 percent in Ghana and 6 percent in Kenya.<sup>6</sup>

These results suggest that earnings rise with firm growth. Because rapidly rising labor costs are likely to constrain firm growth, it is important to understand why this is observed. Underlying Velenchik's empirical approach are many economic theories (e.g., rent sharing; Teal 1996), but as seen, these do not appear to be the whole explanation for the size–wage gap. One possibility is that, as in efficiency wage models, large firms pay relatively higher wages to provide workers with incentives to exert effort.<sup>7</sup> Fafchamps and Soderbom (2006) derive a theoretical model in which the size–wage relationship is generated by the rising cost of monitoring workers as firm size increases, and they estimate the structural parameters using worker–firm-matched data across 10 African countries. The implications of their findings are discussed in the following section.

Although it may not be entirely clear why there is a size–wage gap in the data, there is fairly strong evidence that the gap is not simply an artifact of omitting unobserved skills from the regressions. There is also good evidence that the size–wage gap is quantitatively large. In the case of Ghana, for instance, Soderbom and Teal (2004) estimate that if a firm with 350 employees faced the same wage as a firm with 20 employees, total costs in the large firm would fall by 20–25 percent. One implication of the large size–wage gap is that firm growth is associated with rising labor costs, which, everything else held constant, will tend to erode firms' profits.

So how can growing firms remain profitable? One way would be through improved performance in the form of higher productivity. Soderbom and Teal (2004) look for evidence of increasing returns to scale in Ghanaian manufacturing. Thus as firms grow, higher average labor costs would be offset by higher input productivity. Soderbom and Teal report production function estimates that are robust to unobserved time-invariant heterogeneity across firms (fixed effects) and endogeneity in the factor inputs. They find only weak evidence for increasing returns, and constant returns to scale is not rejected. They argue that the reason large firms can remain profitable is that they face lower capital costs than small firms. One possible reason is better access to formal credit.

**Infrastructure.** Based on several studies of growth determinants in Africa, Collier and Gunning (1999) argue that poor infrastructure is a serious constraint to growth on the continent. Compared with other regions, public expenditure as a share of gross domestic product (GDP) has been higher in Africa, whereas service provision has been worse. The poor infrastructure in Africa is likely to be a particularly severe constraint to manufacturing growth. Many firms spend their own resources directly to buy infrastructure services or provide it on their own. One reason for the prevalence of small manufacturing firms in Africa (see *Enterprise Organization* below) is that transport infrastructure is poorly developed, which creates pockets of demand that tend to generate small-scale localized producers. To be able to take advantage of

international trading opportunities and to engage actively in the emerging system of outsourcing, firms require reliable and cheap infrastructure.

The firm surveys have generated information about how firms perceive infrastructure problems and recently also about the time and monetary costs of different bottlenecks. Bigsten and others (1999a) show that attempts to use data on perceptions to explain productivity is problematic, however. Firms that rank infrastructure problems as very severe are the most productive ones, arguably because these firms sell on larger markets and are therefore more dependent on infrastructure. The use of objective measures (e.g., days in customs, days without telephone connections, days without electricity) is a more promising approach, but so far little analysis is available on the impact of infrastructure on manufacturing firm productivity.<sup>8</sup>

### *Enterprise Organization and the Business Environment*

*Enterprise organization.* A conspicuous characteristic of Africa's manufacturing sector is the preponderance of very small and informal firms, which operate side by side with few large factories. The impact on the efficiency of resource allocation is pertinent. Bigsten, Kimuyu, and Lundvall (2004) investigate whether there are productivity differentials between formal and informal firms in Kenya, where the bulk of manufacturing employment is in the informal sector. Using a production function approach, these researchers find no significant productivity difference between small informal and small formal African-owned firms. A reallocation of firms from the informal to the formal sector would thus not necessarily affect aggregate productivity. Nevertheless, there is little investment and little exporting in the informal sector, and so growth in this sector is unlikely to be a source of significant modernization. Furthermore, wages in the informal sector are low, and contributions to tax revenues miniscule. There is therefore a case for policies to encourage the formalization of informal firms.

In view of this need, why small informal firms do not grow is an important question. Bigsten, Kimuyu, and Lundvall (2004) note that formal firms in Kenya have experienced significant problems in dealing with the authorities relating to corruption, regulatory red tape, and lack of security. In a similar vein Sleuwaegen and Goedhuys (2002) report that the proportion of firms that perceive regulations as an obstacle to growth is lower in the informal than in the formal manufacturing sector in Côte d'Ivoire (the opposite is true for market conditions). With cost increases and no productivity gains from becoming formal, it may thus be rational for African entrepreneurs to choose to start or remain informal. To see a shift in the size distribution from small to medium and large firms, the incentives probably need to change. This topic is discussed again later.

Fafchamps and Soderbom (2006) offer a different perspective on organization, using a principal agent framework to analyze the consequences of inadequate incentives for

workers. A key assumption in their model is that the costs of monitoring workers rise with firm size. As a result of losing control of workers as firm size increases, firms need to motivate workers by raising their wages instead. The theoretical model consists of a system of equations in which the wages of supervisors and workers are determined simultaneously with the number of supervisors and workers employed by the firm. That is, both firm size and wages are treated as endogenous variables. Estimates of the parameters of the structural model indicate that a doubling in the number of production workers will increase the labor cost per unit of effort by 9 percent for Morocco and 11–14 percent for sub-Saharan Africa—the penalty large firms incur to motivate workers. Thus the incentive problem appears to be more severe in sub-Saharan Africa than in Morocco.

Fafchamps and Soderbom (2006) further find that an increase in total factor productivity leads to an increase in worker effort in Morocco but a decrease in sub-Saharan Africa. Firms in sub-Saharan Africa with high total factor productivity hire fewer workers and supervisors (and produce less output) than they would have done had the incentive structure been similar to that in Morocco. Quantitatively, this effect on output is large: a 1 percent increase in total factor productivity raises output 2.9 percent in Morocco but only 1.3 percent in sub-Saharan Africa. If Fafchamps and Soderbom are right in arguing that this is because a firm with high total factor productivity finds it more difficult in sub-Saharan Africa than in Morocco to manage and supervise its labor force, the implications for firm growth are potentially quite significant.

*Business environment.* The business environment also affects how firms do business. Widespread market failures imply that firms face many "holes" in important markets, such as those for insurance and credit. This is especially pronounced for the smallest firms. Fafchamps (2001) identifies two types of responses that aim to reduce the resulting transaction costs: developing relationships and sharing information within networks. Essentially, when search and verification costs are high, it makes sense to establish long-term relationships and share information with other market participants to economize on such costs. The enterprise data have been extensively used to shed light on these mechanisms.

Trade credit is a good example of how supplier relationships can fill in some of the holes in the formal credit market. Fisman (2001) argues that trade credit plays an important role in enterprise financing in Africa, mainly by enabling firms to manage raw material inventories more efficiently and by reducing the likelihood of raw material shortages. Trade credit may therefore contribute to higher productive efficiency. Fisman analyzes these issues using the first wave of RPEd data from Côte d'Ivoire, Kenya, Tanzania, Zambia, and Zimbabwe. Trade credit is the dominant form of payment to suppliers for about a third of the firms. There is much variation across countries, however. Trade credit is the dominant form of payment for 8 percent

of firms in Tanzania but for 69 percent in Zimbabwe. Trade credit appears to be more important in countries where the average firm size is larger, suggesting that large firms are more likely to use trade credit than small ones. In the formal empirical analysis Fisman finds that firms that use supplier credit have significantly higher capacity utilization than firms that do not and are less likely to have raw material stockouts. Furthermore, he finds that these results are particularly strong among firms that are constrained in their access to formal credit. These results are robust to the potential endogeneity of supplier credit, which may arise if firms use trade credit because they have high capacity utilization or substantial inventories, for instance.

Fisman (2001) suggests that these results imply the possibility of substantial productivity gains from an increase in the availability of supplier credit. He also makes the point, however, that without understanding why some firms have access to credit while others do not, it will be difficult to provide solid advice to policymakers. If lack of access to trade credit is driven by market failures—perhaps lack of public information on creditworthiness or poor contract enforcement—policy measures to fix these problems would probably have the desired results, but if firms do not get trade credit because they are unreliable, pose a big credit risk, or do not want trade credit, then it is not clear that improving access should be a policy priority.

Fisman and Raturi (2004) investigate whether various firm and entrepreneurial characteristics affect the likelihood that firms in Ghana, Kenya, Tanzania, Zambia, and Zimbabwe get trade credit. The main focus of this analysis is to see whether trade credit is more common among firms that purchase inputs from suppliers that face considerable competition. The researchers find this to be the case and interpret this finding as evidence that suppliers operating in a competitive market use credit to create buyer loyalty. Fisman and Raturi argue that this is consistent with a situation in which firms that want trade credit must establish creditworthiness and, because of information imperfections, must do so every time they switch suppliers. This form of market failure creates rents for suppliers and increases costs for manufacturers.<sup>9</sup> A related result, documented by Fafchamps (2000), is that Asian and European entrepreneurs have better access to trade credit than African entrepreneurs. His hypothesis is that this is due to ethnic networking in contract enforcement.

There is still no complete picture of why African firms get trade credit, and this research suggests that strengthening credit market institutions may be important. In the specific context considered by Fisman and Raturi (2004), the policy implication is that establishing efficient credit rating agencies that keep a record of companies' credit histories would erode the de facto monopoly power of suppliers and reduce costs for manufacturers.

### *Enterprise Growth*

From a policy point of view a good understanding of the relationships between growth and certain enterprise characteristics is important, because it can offer guidance

to policymakers about what types of firms are likely to be more successful and better at creating jobs in the future. The relationship between enterprise size and growth is of particular interest for the African manufacturing sector, because most firms in Africa are very small. How realistic is it to hope that some of these firms will grow and become successful large firms? The relationship between firm age and growth is also important. For example, if young firms grow quickly, policy measures aimed at encouraging entry may have significant growth effects in the short and medium term.

A common way of investigating the relationships between growth on the one hand and size and age on the other is to run regressions of the growth rate of employment between two periods on the explanatory variables employment and age in the initial period. Several such studies have found a negative relationship between size and growth (Ramachandran and Shah 1999; Mazumdar and Mazaheri 2003).

Sleuwaegen and Goedhuys (2002) argue that such regressions may be too restrictive by not allowing for nonlinear relationships. When these researchers augment the basic growth model with terms nonlinear in size and age, they find strong evidence of a positive effect on growth of the interaction between age and size. Everything else equal, the relationship between age and growth is less negative (or more positive) for large firms than for small firms, and the relationship between size and growth is less negative (or more positive) for old firms than for young ones. Thus, a firm that is small when it enters the market grows relatively fast initially, but growth rates slows over time and eventually the firm may start to contract. For a large firm entering the market, the pattern is the opposite. Growth rates are low initially—in fact, employment may fall rapidly in the initial years of operation—but eventually the favorable age effect (driven by the positive interaction effect between size and age) kicks in, and growth rates increase. The implication of the results is that small and large firms have very different growth patterns: high growth tends to be observed mostly among the small and young firms and the large and old firms. These results square with descriptive statistics indicating that few small firms ever grow up to become large (Biggs, Ramachandran, and Shah 1999; Biggs and Shah 2003).

Although interesting, these results should probably be interpreted with a dose of caution because it is possible that the estimated correlation between size and growth is spurious. The problem, dubbed "the Regression Fallacy" by Davis, Haltiwanger, and Schuh (1996), arises whenever there are transitory fluctuations in size or transitory measurement errors in observed size. The resulting bias in the estimated relationship between initial size and growth is negative, so failure to address this problem can produce a picture of the growth of small firms that is too good. The researchers propose that one way of dealing with the problem is to use average size over the observed period rather than initial size as the measure of enterprise size. Adopting this procedure, Mazumdar and Mazaheri (2003) present descriptive statistics

indicating at most a weak negative relationship between size and growth in Cameroon, Côte d'Ivoire, Ghana, Kenya, Tanzania, Zambia, and Zimbabwe.

Enterprise characteristics other than size and age have been shown to matter for growth as well. Ramachandran and Shah (1999) find that firms in Kenya, Tanzania, Zambia, and Zimbabwe that are controlled by minority entrepreneurs (Asian or European) did better than those controlled by indigenous entrepreneurs.<sup>10</sup> The coefficients on secondary and university education are highly significant and positive for African entrepreneurs, whereas none of the entrepreneurial variables was significant in the non-African model. The researchers hypothesize that non-Africans benefit from various advantages of being a minority, such as access to informal networks, credit, and informal contractual mechanisms, and that for African entrepreneurs education may substitute for access to networks. Mengistae (1999) uses data for Ethiopian manufacturing firms to show that there also are significant differences between indigenous ethnic groups in terms of entrepreneurial success. Related to firm growth is firm survival. Exit rates in African manufacturing are high, and they are highest among the smallest firms. Frazer (2005) documents strong empirical evidence of a positive association between productivity and the likelihood of firm survival among Ghanaian manufacturing firms. The results imply that a one standard deviation increase in total factor productivity is associated with a 1.8 percentage point fall in the likelihood of firm exit over the subsequent two-year period, a reasonably large effect considering that the average exit rate is 6.3 percent. A related study by Soderbom, Teal, and Harding (2006) based on pooled panel data on firms in Ghana, Kenya, and Tanzania over a five-year period finds that total factor productivity affects firm survival among large firms but not small ones. In other words, being relatively more productive does not prevent firms from going out of business if they are small. Among larger firms, however, exit rates do depend on productivity, which is consistent with a survival of the fittest process, as documented in the United States, the United Kingdom, and many middle income countries. This result is also consistent with the findings reported by Frazer (2005).

### *Investment*

This section reviews the empirical literature on investment in Africa's manufacturing sector, focusing on financial constraints and risk. African financial markets are the least developed in the world, and development economists have long held the view that this impedes growth. For investment the main problem is that firms with profitable investment projects often cannot use external funds to finance such projects.

That there are financial imperfections in Africa can hardly be disputed, but such imperfections translate into binding constraints only if firms have a desire to invest. Bigsten and others (1999b) is one of the first studies analyzing whether investment

among African firms is hampered by lack of external finance. That study follows a fairly conventional approach in the empirical investment literature, testing whether investment is sensitive to changes in cash flow among firms observed in the early and mid-1990s in Cameroon, Ghana, Kenya, and Zimbabwe. The evidence indicates that there is a statistically significant profit effect on investment, which suggests the presence of credit constraints. With point estimates on the profit term ranging between 0.06 and 0.10, the magnitude of the effect is small, however: only between \$0.06 and \$0.10 cents of an additional \$1 earned in profits are invested.

Subsequent research based on RPED data confirms that investment is not particularly sensitive to changes in profits. In an in-depth analysis of the manufacturing sector in Zimbabwe 1992–94, Fafchamps and Oostendorp (1999) show that the sensitivity of investment to cash flow is low, even among small firms. Soderbom (2002) obtains a similar result for Kenya. Mazumdar and Mazaheri (2003) use a sample of six countries [adding Côte d'Ivoire and Zambia to the set of countries considered by Bigsten and others (1999b)] and report an estimated profit coefficient of 0.08, which is very similar to that of Bigsten and others (1999b). Mazumdar and Mazaheri also split the sample by firm size and obtain a profit coefficient of 0.09 in the subsample of small firms. They interpret the larger profit coefficient among small firms as evidence of greater credit constraints among small firms than large ones, which seems likely. Reinikka and Svensson (2001) obtain a profit coefficient of 0.08 based on a sample of Ugandan manufacturing firms for 1996–97. These researchers too find a larger profit effect among smaller firms, which is consistent with the notion that credit access is more of a problem for small firms. However, the point estimate of the coefficient on profits among small firms is 0.11, and so quantitatively the effect is not particularly large even for small firms.

The second-hand market for used machinery in Africa is shallow at best, and so investment expenditures are often "sunk" (investment is irreversible). Theories of irreversible investment under uncertainty predict that investment will be slower to respond to demand shocks if uncertainty is high. As mentioned, testing empirically for the effects of risk is difficult, as risk is not easily measured. The design of innovative survey questions has been a major step forward. Pattillo (1998) uses panel data on Ghanaian manufacturing firms from 1994 and 1995 to test hypotheses from models of irreversible investment under uncertainty. Based on data on entrepreneurs' subjective (or perceived) probability distribution over future demand, Pattillo calculates the variance of demand and uses this as the measure of uncertainty. Empirical results indicate that uncertainty has a negative effect on investment, an effect that is more pronounced for firms with more irreversible investment. Darku (2001) undertakes a similar inquiry based on firm-level data from a survey in Uganda in 1998. Measuring uncertainty in the same way as Pattillo (1998), Darku finds a negative relation between uncertainty and investment and documents that this effect is stronger among firms with more irreversible

investment. These studies provide direct evidence of a negative effect of perceived uncertainty on investment.

High uncertainty results in a high risk premium in the required return on invested capital, suggesting that African manufacturing firms have high opportunity costs of capital. Bigsten and others (1999b) argue that this is indeed the case, reporting much higher average returns to capital than among firms in more developed countries (table 3). There is a striking similarity in average investment rates across both the African and the European countries considered, but the rates of return on capital are much higher in the African countries. The researchers infer from this that the cost of capital is relatively high in Africa and maintain that this is consistent with a negative effect of uncertainty on investment. Fafchamps and Oostendorp (1999) take a similar view by arguing that uncertainty is a plausible explanation for why investment remained low in Zimbabwe.

Using the same data set as in the 1999 study, but adding Zambia, Bigsten and others (2005) examine whether investment is affected by irreversibility and fixed adjustment costs. If so, one would expect to see a significant share of zero investments. Combining the information presented by Bigsten and others (2005) and that of Reinikka and Svensson (2001) on Uganda, table 4 summarizes the proportions of nonzero investments in this data set by country and firm size. With the exception of Uganda and Zimbabwe, the proportion of positive investment is lower than 0.5 across all size categories, meaning that most firms in these categories do not invest at all during a typical year. There is a weak positive relationship between firm size and propensity to invest, although among the largest firms in Ghana only one in five firms invest in a representative year. Investment activity is generally higher in Zimbabwe than in the other countries, although a third of firms in the Zimbabwean sample still

**Table 3.** Investment Rates and Returns on Fixed Capital in Selected African and European Countries (Percent)

	<i>Investment rate</i>	<i>Average return on fixed capital</i>
Africa		
Cameroon	0.11	1.36
Ghana	0.13	3.63
Kenya	0.11	1.82
Uganda	0.12	0.75
Zimbabwe	0.12	0.85
Europe		
Belgium	0.13	0.18
France	0.11	0.12
Germany	0.12	0.16
United Kingdom	0.12	0.13

*Source:* All countries except Uganda (Bigsten and others 1999b); Ugandan (Reinikka and Svensson 2001).

**Table 4.** Proportions of Nonzero Investments Among Firms in Six African Countries, by Firm Size

<i>Number of employees</i>	<i>Cameroon</i>	<i>Ghana</i>	<i>Kenya</i>	<i>Uganda</i>	<i>Zambia</i>	<i>Zimbabwe</i>
1–5	0.21	0.31	0.44	n.a.	0.29	0.53
6–20	0.29	0.44	0.40	n.a.	0.29	0.51
21–100	0.24	0.48	0.41	n.a.	0.28	0.63
101 or more	0.38	0.20	0.44	n.a.	0.38	0.71
Total	0.29	0.32	0.42	0.52	0.31	0.66

*Source:* Bigsten and others (2005) and Reinikka and Svensson (2001).

refrain from investing in a typical year. The researchers also show that investment is "lumpy"; whenever firms do invest, they invest a lot. This suggests that fixed adjustment costs may be important. However, more formal analyses of the decision to invest using a dynamic discrete choice model point to irreversibility as the main explanation for low investment in Africa. Reducing uncertainty, or improving the market for second-hand fixed capital, is therefore likely to positively affect investment.

### *Exports*

As mentioned, manufacturing firms in Africa operate in small domestic markets. To expand production, firms may have to orient part of their production toward exporting. What factors prevent African firms from entering export markets? And are there any benefits, other than market enlargement, associated with exporting? In particular, is there any evidence that firms become more productive as a result of exporting, perhaps because of contacts with foreign customers or exposure to international competition?

At least two key factors determine whether a firm participates in the export market: the level of entry barriers and the firm's cost efficiency. In an influential study of the decision to export, Roberts and Tybout (1997) argue that entering the export market for the first time may be associated with a fixed entry cost. For instance, the firm may need to set up a marketing department to investigate marketing channels and meet export orders. Although entry costs are typically not observed, Roberts and Tybout argue that state dependence in exports—that lagged participation affects the likelihood of current participation—may constitute indirect evidence of entry costs. Thus in the absence of entry costs, firms will switch in and out of the export market independently of whether they have exported in the past. If there are significant entry costs, however, firms that have incurred these costs in the past (and thus will not have to incur them again) will be more likely to export in subsequent periods than firms that have not, because exporting is less costly for insiders than for outsiders.

In testing for state dependence in exports, it is important to recognize the possibility that lagged exporting status is not strictly exogenous. In particular, if there is unobserved time-invariant heterogeneity across firms in the propensity to export, failure to control for such heterogeneity will tend to bias the results toward state dependence. Roberts and Tybout deal with this issue, often referred to as the initial conditions problem, using techniques developed by Heckman (1981). In their empirical application, which is based on data on manufacturing plants in Colombia, Roberts and Tybout find strong evidence that insiders are more likely to export than outsiders. This indicates that entering the export market is associated with significant fixed costs.

Bigsten and others (2004) carry out a similar analysis based on RPEd data from the early 1990s for Cameroon, Ghana, Kenya, and Zimbabwe. Controlling for some factors, including firm-specific time-invariant unobserved heterogeneity, they find that past export status has a significant effect on the propensity to export. The magnitude of this effect is large: for the average firm that did some exporting in the previous period, the likelihood of exporting in the current period is about 0.57, whereas the likelihood of exporting for an otherwise identical firm that did not export in the previous period is 0.18. Thus, for a nonexporting firm with the average characteristics, entering the export market raises the probability of exporting in the next period from less than one in five to more than one in two. This is attributed to large entry costs. Van Biesebroeck (2005a) obtains similar results for a larger sample of nine sub-Saharan Africa countries using a similar approach.

The second factor that determines whether a firm will export is its cost efficiency. Clerides, Lach, and Tybout (1998) derive a model in which exporting is associated with a fixed cost that is incurred in every period of exporting. In this model firms with marginal costs below some threshold choose to export, whereas firms with marginal costs above the threshold remain focused on the domestic market. Thus, this model predicts that relatively efficient firms will self-select into the export market. The study by Bigsten and others (2004) looks for evidence of self-selection in export markets and finds this effect to be relatively weak. Van Biesebroeck (2005a) finds a somewhat stronger self-selection effect in a larger sample. Both studies report results suggesting that causality runs in the other direction, from exporting to efficiency.

The fact that exporters tend to be more efficient than nonexporters is a common result in both rich and poor countries. Mengistae and Pattillo (2004) report a positive correlation between productivity and exporting among firms in Ethiopia, Ghana, and Kenya. Whether this is because exporting actually causes efficiency gains has received considerable attention in the literature recently. From a policy perspective whether firms in developing countries learn from exporting is an important issue because learning by exporting can reduce the competitiveness gap endogenously through increased international trade. One methodological challenge in

testing for learning by exporting is that exporting will be endogenous if efficient firms self-select into the export market: efficiency and exports may be correlated even in the absence of learning effects.

Clerides, Lach, and Tybout (1998) develop an econometric framework for teasing out the relative importance of learning effects and self-selection effects. Key features of this approach are that exports are determined by efficiency and that there is a common unobserved time-invariant factor that affects both the propensity to export and firm-level efficiency. Using the econometric framework proposed by Clerides, Lach, and Tybout (1998), Bigsten and others (2004) find relatively strong evidence of learning effects in that participating in export markets has positive effects on total factor productivity in subsequent periods. Van Biesebroeck (2005a) uses similar methods and confirms the presence of learning effects. The quantitative effect of exporting appears to be large. The estimates reported by Bigsten and others imply that exporting is associated with a productivity gain of 7–8 percent in an output production function, which corresponds to productivity gains in terms of value added of 20–25 percent in the short run and up to 50 percent in the long run. These estimates are in line with the results reported by Van Biesebroeck (2005a).

## Conclusions, Future Research, and Some 'Thoughts on Policy

A decade of surveys in Africa has greatly improved our understanding of the factors that drive the choices and outcomes in the manufacturing sector. Four main findings emerge from the research.

First, investment in physical capital has remained low, more because of uncertainty than because of a severe credit constraint. There is some evidence that lack of credit has been a problem for small firms, but although the profit effect on investment is larger for small than for large firms (Reinikka and Svensson 2001), it is still quite small. Analysis of firms' borrowing behavior paints a similar picture: on average the desire for formal credit has been relatively modest, although demand for credit is relatively high among very small firms. The most likely explanation for why a lack of credit has not been a major factor in explaining the low levels of investment over the last decade is that few firms could identify strong investment opportunities during this period. This does not mean that the financial sector reforms implemented in many African countries in the 1990s were unnecessary but only that the constraints that were relaxed were not binding at the time. When firms expand and the need for formal borrowing increases, the financial reforms may turn out to have a higher payoff.

Second, exports have remained low throughout the period, and research indicates that the high costs of entering the export market may be part of the reason. This has two potentially important policy implications. First, if incentives can be created for

firms to enter the export market, firms are likely to remain in the export market for some time. Second, high entry costs imply that there is a large set of firms that remain focused on the domestic market even though they are internationally competitive. Reducing entry costs will give these firms access to a larger market.

Third, evidence is fairly strong that exporting leads to efficiency gains—there is learning by exporting. Studies looking at this issue in other regions tend not to obtain this result (Clerides, Lach, and Tybout 1998; see Blalock and Gertler 2004 for a counter example). One possible explanation is that the potential gains from exporting are large in Africa because of high trade restrictions in the past and a large technological gap with developed countries. Exporting thus offers maximum scope for the increased discipline of competition, and contact with foreign customers provides maximum scope for learning opportunities. Arguably, if exporting induces efficiency in any environment, it should do so in Africa. This implies that Africa may have much to gain from orienting its manufacturing sector toward exporting.

Fourth, there is strong evidence that earnings vary across individuals with similar skills and thus that labor costs vary across firms even though the quality of labor may be similar. Wages are higher in large firms than in small, but this can be attributed only partly to differences in worker characteristics. Why this is so is not fully understood, but the effect is quantitatively large.

Improving our understanding of why these results are observed seems an important area for future research. Other issues also need further research. Although much has been learned about enterprise growth from the enterprise data, there is still uncertainty about the association between enterprise size and growth. It may be that growth rates are independent of enterprise size, in which case policymakers should not expect systematically higher growth rates in certain size segments of the sector. To address some of the methodological challenges discussed, researchers will need to use data covering firms over a relatively long period. Access to panel data with a reasonably long time dimension is therefore important. Furthermore, whether the fact that small and relatively productive firms have high exit rates (Soderbom, Teal, and Harding 2006) is a cause for concern depends on why this result is observed. It would be socially wasteful if exits were involuntary, because of some uninsurable and temporary shock. However, if exit is voluntary, the result of the entrepreneur finding a more profitable occupation elsewhere, it is less clear that this is a problem.

The research results based on the African enterprise data have implications for policy. More jobs can be created only if the industrial sector expands. With domestic markets still very small, most expansion will probably have to be through exports. The survey data indicate that the rewards from participation in international markets may be substantial but also that firms are reluctant to enter the export market because of high costs. Facilitating exporting would thus appear an important part of the policy package. With poor infrastructure and a weak regulatory system, Africa is

at a disadvantage, so reforms in these areas would seem important. That good policies can make a difference is clear from Madagascar, for instance, where the creation of an efficient export processing zone led to a very rapid expansion in the textiles and garment sector in the 1990s.

Furthermore, to be internationally competitive, more investment is needed.<sup>12</sup> Because the survey data support the notion that uncertainty hampers investment, sound and credible policies consistently pursued over time should reduce the risk perceived by entrepreneurs. Governance is likely to be central, because without good governance, uncertainty will persist and investors' response to other initiatives will be weak. Stringing together several years of good governance and good policies can be hard in a situation where people do not see quick results, and resisting policy reversals and populist policies will therefore be a challenge.

In addition, scarce resources — land, skills, physical capital, and raw material — need to be channeled to firms that are productive and in a good position to export. An important question is what to do with the large informal manufacturing sector in Africa. Based on Kenyan enterprise data, Bigsten, Kimuyu, and Lundvall (2004) argue that the informal sector is large primarily because being formal is costly. Thus, addressing the basic governance and investment climate issues would help encourage entrepreneurs to move from the informal to the formal sector.

Of course, that shift will not occur over night, and during a transition period steps could be taken to improve performance within the informal sector. Enterprise data show that some of the key economic fundamentals, such as total factor productivity, are reasonably good in informal firms (Bigsten, Kimuyu, and Lundvall 2004). However, vocational training and other capacity-building efforts have so far been directed mainly to the formal sector while largely neglecting the needs of informal firms. One likely welfare loss stems from the lack of access to credit, and hence forgone investment, for informal firms, because of their ambiguous legal status and lack of proper accounts, as well as the lack of credit-rating procedures relevant to informal firms. Because of the small absolute size of the loans required by informal firms, it should be feasible to support alternative procedures for the provision of credit.

Although most African firms have experienced limited success at best, there is also a class of firms that has performed very well. This general finding suggests that African manufacturing firms have the potential to perform well. To enable firms to improve their performance, key goals for African industrial policy should be to facilitate the adoption of new technology and the penetration of new markets. Countries that cannot break out of the current situation — in which most manufacturing firms focus on supplying the domestic market with low value-added products — are unlikely to see a significant expansion of jobs in the manufacturing sector or to have manufacturing play a major role in reducing poverty.

## Notes

Arne Bigsten is a professor of economics in the Department of Economics at Goteborg University; his e-mail address is arne.bigsten@economics.gu.se. Måns Söderbom is a research officer at the Centre for the Study of African Economies in the Department of Economics at the University of Oxford; his e-mail address is mans.soderbom@economics.ox.ac.uk. The authors thank Shantayanan Devarajan, three anonymous referees, John Nasir, Adrian Wood, and participants at seminars at Goteborg University and Oxford University for helpful comments on earlier drafts of the article. They are thankful to the World Bank for funding the research.

1. This is not the first study to take stock of what can be learned from data on manufacturing firms in Africa. Tybout (2000) reviews the research on firm performance in developing regions, including Africa. Mazumdar and Mazaheri (2003) focus exclusively on African manufacturing. Collier and Gunning (1999) discuss the problems faced by Africa's manufacturing sector as part of their survey paper on Africa's general economic problems. All of these studies cover data on African firms over the period 1991–95, the period covered by the first wave of data from the RPED. More recently, Kingdon, Sandefur, and Teal (2004) and Fafchamps and Soderbom (2004) have provided overviews of the conditions for job creation and labor demand in Africa, drawing on the firm surveys among other sources. This review is based on more recent data than that of Collier and Gunning, Tybout, and Mazumdar and Mazaheri and in looking at a wide range of results that have come out of the research on the firm data it has a broader focus than the reviews by Kingdon, Sandefur, and Teal and Fafchamps and Soderbom.

2. For instance, Bigsten and others (1999b) analyze data on manufacturing firms in four African countries and report that profit to capital ratios are more than four times lower than the median for one in four firms and more than four times higher than the median in one in four firms (and this is not a result primarily of cross-country differences). A similar finding of large variability across firms holds for most of the key performance variables in the data (productivity, export intensity, and investment).

3. Investment climate is a very broad concept. Its key components are the institutional, policy, and regulatory environment in which firms operate (Dollar, Hallward-Driemeier, and Mengistae 2005). Examples of recently studied dimensions of the investment climate include the quality of infrastructure, the nature of business regulations and their enforcement, the prevalence of credit constraints, the quality of governance, general conditions for private investment and enterprise growth, economic freedom, country credit ratings, human development, environmental sustainability, and civil rights.

4. See <http://www.csae.ox.ac.uk/datasets/main.html>.

5. Enterprise surveys are not the only useful data for analyzing the role of the investment climate. In fact, because many aspects of the investment climate are constant across firms within an economy, it will sometimes be better to measure the key dimensions of the investment climate at the country level—perhaps by means of a few case studies—and then to conduct the empirical analysis at the aggregate level. One major project based on such a procedure has generated the World Bank and International Finance Corporation Doing Business Database (World Bank 2005), which contains data on aspects of the regulatory environment facing private firms in a large number of countries and economic regions—for example, how long it takes to start a business, how hard it is to hire and fire workers, how well contracts are enforced, and so on. The construction of these data is based on case studies of laws and regulations in each country and does not involve surveying individual firms.

6. Ghana:  $\exp(0.15\ln(40) - 0.15\ln(20)) - 1 = 0.11$ . Kenya:  $\exp(0.08\ln(40) - 0.08\ln(20)) - 1 = 0.06$ .

7. The idea that efficiency wages drive the wage–size gap is considered by Velenchik (1997), but not tested directly.

8. See Dollar, Hallward-Driemeier, and Mengistae (2005) for such analysis based on firm-level data from Bangladesh, China, Pakistan, and India.

9. There is also evidence that ethnicity plays a role in whether firms use trade credit at all (Fisman 2003). This too may be consistent with information imperfections.

10. In an early study of the relations between minorities and entrepreneurial success, Kilby (1983) argues that minority entrepreneurs often have superior initial endowments of capital, technology, and knowledge of markets and have acquired traditions that help them raise productivity.

11. One of the main objectives of investment is to get access to better technology. Direct transfers of technology through technology contracts and the like to African manufacturing firms have been very limited (see, e.g., Bigsten and Kimuyu 2002, on Kenya).

12. Currently, foreign investors do not appear to see Africa as a promising location for investment, and many Africans share this view and keep a large share of their wealth outside Africa (about 40 percent according to Collier, Hoeffler, and Pattiilo 2001).

## References

- Batra, G., D. Kaufmann, and A. H. W. Stone. 2003. *Investment Climate Around the World: Voices of the Firms from the World Business Environment Survey*. Washington, D.C.: World Bank.
- Biggs, T., and M. Shah. 2003. "The Problem of African Entrepreneurial Development." In G. S. Fields and G. Pfeffermann, eds., *Pathways Out of Poverty: Private Firms and Economic Mobility in Developing Countries*. Boston: Kluwer Academic Publishers.
- Biggs, T., V. Ramachandran, and M. Shah. 1999. "The Determinants of Enterprise Growth in Sub-Saharan Africa: Evidence from the Regional Program for Enterprise Development." RPED Discussion Paper 135. Washington, D.C.: World Bank.
- Bigsten, A., and P. Kimuyu, eds. 2002. *Structure and Performance of Manufacturing in Kenya*. London: Palgrave Macmillan.
- Bigsten, A., P. Kimuyu, and K. Lundvall. 2004. "What to Do with the Informal Sector?" *Development Policy Review* 22(6):701–15.
- Bigsten, A., P. Collier, S. Dercon, M. Fafchamps, B. Gauthier, J. W. Gunning, A. Oduro, R. Oostendorp, C. Pattillo, M. Soderbom, F. Teal, and A. Zeufack. 1999a. "Economic Infrastructure and Productivity in Sub-Saharan Africa: A Firm-Level Analysis." Goteborg, Sweden: Goteborg University.
- Bigsten, A., P. Collier, S. Dercon, B. Gauthier, J. W. Gunning, A. Isaksson, A. Oduro, R. Oostendorp, C. Pattillo, M. Soderbom, M. Sylvain, F. Teal, and A. Zeufack. 1999b. "Investment in Africa's Manufacturing Sector: A Four Country Panel Data Analysis." *Oxford Bulletin of Economics and Statistics* 61(4):489–512.
- Bigsten, A., P. Collier, S. Dercon, M. Fafchamps, B. Gauthier, J. W. Gunning, A. Oduro, R. Oostendorp, C. Pattillo, M. Soderbom, F. Teal, and A. Zeufack. 2003. "Credit Constraints in Manufacturing Enterprises in Africa." *Journal of African Economies* 12(1):104–25.
- \_\_\_\_\_. 2004. "Do African Manufacturing Firms Learn from Exporting?" *Journal of Development Studies* 40(3):115–71.
- Bigsten, A., P. Collier, S. Dercon, M. Fafchamps, B. Gauthier, J. W. Gunning, R. Oostendorp, C. Pattillo, M. Soderbom, and F. Teal. 2005. "Adjustment Costs and Irreversibility as Determinants of Investment: Evidence from African Manufacturing." *Contributions to Economic Analysis & Policy* 4(1):1–27.
- Blalock, G., and P. J. Gertler. 2004. "Learning from Exporting Revisited in a Less Developed Setting." *Journal of Development Economics* 75(2):397–416.
- Clerides, S., S. Lach, and J. Tybout. 1998. "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico and Morocco." *Quarterly Journal of Economics* 113(3):903–48.
- Collier, P. 2000. "Africa's Comparative Advantage." In H. Jaliian, M. Tribe, and J. Weiss, eds., *Industrial Development and Policy in Africa*. Cheltenham, UK: Edward Elgar.
- Collier, P., and J. W. Gunning. 1999. "Explaining African Economic Performance." *Journal of Economic Literature* 37(1):64–111.

- Collier, P., A. Hoeffler, and C. Pattillo. 2001. "Flight Capital as a Portfolio Choice." *World Bank Economic Review* 15(1):55–80.
- Darku, A. B. 2001. "Private Investment, Uncertainty, and Irreversibility in Uganda." *African Finance Journal* 3(1):1–25.
- Davis, S. J., J. C. Haltiwanger, and S. Schuh. 1996. *Job Destruction and Destruction*. Cambridge, Mass.: The MIT Press.
- Dollar, D., M. Hallward-Driemeier, and T. Mengistae. 2005. "Investment Climate and Firm Performance in Developing Economies." *Economic Development and Cultural Change* 54(1):1–32.
- Eifert, B., A. Gelb, and V. Ramachandran. 2005. "Business Environment and Comparative Advantage in Africa: Evidence from the Investment Climate Data." RPEID Report 126. Washington, D.C.: World Bank.
- Fafchamps, M. 2000. "Ethnicity and Credit in African Manufacturing." *Journal of Development Economics* 61(1):205–35.
- . 2001. "Networks, Communities, and Markets in Sub-Saharan Africa: Implications for Firm Growth and Investment." *Journal of African Economies* 10(Suppl.2):109–42.
- Fafchamps, M., and R. Oostendorp. 1999. "Investment." In J. W. Gunning and R. Oostendorp, eds., *Industrial Change in Africa: Micro Evidence on Zimbabwean Firms Under Structural Adjustment*. London: Macmillan.
- Fafchamps, M., and M. Soderbom. 2004. "Climate for Job Creation." In *Employment Issues—Regional Stocktaking Review*. Washington, D.C.: World Bank.
- . 2006. "Wages and Labor Management in African Manufacturing." *Journal of Human Resources* 41(2):346–79.
- Fafchamps, M., J. W. Gunning, and R. Oostendorp. 2000. "Inventories and Risk in African Manufacturing." *Economic Journal* 110(466):861–93.
- Fisman, R. 2001. "Trade Credit and Productive Efficiency in Developing Countries." *World Development* 29(2):311–21.
- . 2003. "Ethnic Ties and the Provision of Credit: Relationship-Level Evidence from African Firms." *Advances in Economic Analysis and Policy* 3(1):1–18.
- Fisman, R., and M. Raturi. 2004. "Does Competition Encourage Credit Provision? Evidence from African Trade Credit Relationships." *Review of Economics & Statistics* 86(1):345–52.
- Frazer, G. 2005. "Which Firms Die? A Look at Manufacturing Firm Exit in Ghana." *Economic Development and Cultural Change* 53(3):585–617.
- Gunning, J. W., and T. Mengistae. 2001. "Determinants of African Manufacturing Investment: The Microeconomic Evidence." *Journal of African Economies* 10(Suppl.2):48–80.
- Habyarimana, J. 2003. *The Benefits of Banking Relationships: Evidence from Uganda's Banking Crisis*. Cambridge, Mass.: Harvard University.
- Heckman, J. J. 1981. "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process." In C. F. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, Mass.: The MIT Press.
- Jones, P. 2001. "Are Educated Workers Really More Productive?" *Journal of Development Economics* 64(1):57–79.
- Kiby, P. 1983. "The Role of Alien Entrepreneurs in Economic Development." *American Economic Review* 73(2):107–11.
- Kingdon, G., J. Sandefur, and F. Teal. 2004. "Patterns of Labor Demand." In *Employment Issues—Regional Stocktaking Review*. Washington, D.C.: World Bank.

- Mazumdar, D., and A. Mazaheri. 2003. *The African Manufacturing Firm*. London: Routledge.
- Mengistae, T. 1999. "Indigenous Ethnicity and Entrepreneurial Success in Africa: Some Evidence from Ethiopia." Policy Research Working Paper 2534. Washington, D.C.: World Bank.
- Mengistae, T., and C. Pattillo. 2004. "Export Orientation and Productivity in Sub-Saharan Africa." *IMF Staff Papers* 51(2):327–53.
- Pattillo, C. 1998. "Investment, Uncertainty, and Irreversibility in Ghana." *IMF Staff Papers* 45(3):522–53.
- Pattillo, C., and M. Soderbom. 2001. *Do Risk Averse Entrepreneurs Get Lower Profits?* Oxford: University of Oxford.
- Psacharopoulos, G. 1994. "Returns to Investment in Education: A Global Update." *World Development* 22(9):1325–43.
- Psacharopoulos, G., and H. A. Patrinos. 2002. "Returns to Investment in Education: A Further Update." Policy Research Working Paper 2881. Washington, D.C.: World Bank.
- Ramachandran, V., and M. K. Shah. 1999. "Minority Entrepreneurs and Firm Performance in Sub-Saharan Africa." *Journal of Development Studies* 36(2):71–87.
- Reinikka, R., and J. Svensson. 2001. "Confronting Competition: Investment, Profit, and Risk." In R. Reinikka and P. Collier, eds., *Uganda's Recovery*. Washington, D.C.: World Bank.
- Roberts, M., and J. Tybout. 1997. "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs." *American Economic Review* 87(4):545–64.
- Sleuwaegen, L., and M. Goedhuys. 2002. "Growth of Firms in Developing Countries, Evidence from Côte d'Ivoire." *Journal of Development Economics* 68(1):117–35.
- Soderbom, M. 2002. "Investment Behaviour." In A. Bigsten and P. Kimuyu, eds., *Structure and Performance of Manufacturing in Kenya*. London: Palgrave Macmillan.
- Soderbom, M., and F. Teal. 2004. "Size and Efficiency in African Manufacturing Firms: Evidence from Firm-Level Panel Data." *Journal of Development Economics* 73(1):369–94.
- Soderbom, M., F. Teal, and A. Harding. 2006. "The Determinants of Survival Among African Manufacturing Firms." *Economic Development and Cultural Change* 54(3):533–56.
- Soderbom, M., F. Teal, and A. Wambugu. 2005. "Unobserved Heterogeneity and the Relation Between Earnings and Firm Size: Evidence from Two Developing Countries." *Economics Letters* 87(2):153–9.
- Soderbom, M., F. Teal, A. Wambugu, and G. Kahyarara. 2006. "The Dynamics of Returns to Education in Kenyan and Tanzanian Manufacturing." *Oxford Bulletin of Economics and Statistics* 68(3):261–88.
- Teal, F. 1996. "The Size and Sources of Economic Rents in a Developing Country Manufacturing Labour Market." *Economic Journal* 106(437):963–76.
- Tybout, J. 2000. "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?" *Journal of Economic Literature* 38(1):11–44.
- Van Biesebroeck, J. 2005a. "Exporting Raises Productivity in Sub-Saharan African Manufacturing Firms." *Journal of International Economics* 67(2):373–91.
- . 2005b. "Firm Size Matters: Growth and Productivity Growth in African Manufacturing." *Economic Development and Cultural Change* 53(3):545–83.
- Velenchik, A. D. 1997. "Government Intervention, Efficiency Wages and the Employer Size Effect in Zimbabwe." *Journal of Development Economics* 53(2):305–38.
- World Bank. 2004. *World Development Report 2005: A Better Investment Climate for Everyone*. New York: Oxford University Press.
- . 2005. *Doing Business in 2005: Removing Obstacles to Growth*. Washington, D.C.: World Bank, International Finance Corporation, and Oxford University Press.



---

# Is There a Case for Industrial Policy? A Critical Survey

---

Howard Pack • Kamal Saggi

---

*What are the underlying rationales for industrial policy? Does empirical evidence support the use of industrial policy for correcting market failures that plague the process of industrialization? This article addresses these questions through a critical survey of the analytical literature on industrial policy. It also reviews some recent industry successes and argues that public interventions have played only a limited role. Moreover, the recent ascendance and dominance of international production networks in the sectors in which developing countries once had considerable success implies a further limitation on the potential role of industrial policies as traditionally understood. Overall, there appears to be little empirical support for an activist government policy even though market failures exist that can, in principle, justify the use of industrial policy.*

Many countries in recent years have encountered great disappointment with the results of pursuing the conventional economic policies that John Williamson (1990) crystallized and named the Washington Consensus. Although few countries ever followed the pristine form of this consensus, some countries in East Asia adhered to many (but hardly all) of its components and experienced extraordinarily rapid growth for three decades or more. Although there was a brief and sharp recession in some of these countries during the 1997–99 financial crisis, most have rebounded, with the exception of Indonesia. Yet other countries that have gotten their macroeconomic and trade regimes much closer to the idealized consensus than the Asian countries did have failed to experience comparable growth. In many Latin American countries and in some African countries, there is an understandable search for a magic solution, and many policymakers have expressed interest in some form of industrial policy.

Few phrases elicit such strong reactions from economists and policymakers as *industrial policy*. As Evenett (2003) notes, industrial policy means different things to different people. This article defines industrial policy as any type of selective

government intervention or policy that attempts to alter the structure of production in favor of sectors that are expected to offer better prospects for economic growth in a way that would not occur in the absence of such intervention in the market equilibrium. It is not surprising that those who believe strongly in the efficient working of markets view any argument in favor of industrial policy as fiction or, worse, an invitation to rent seeking, whereas those who believe that market failures are pervasive think that economic development requires a liberal dose of industrial policy.

This article addresses arguments for and against industrial policy and asks whether empirical evidence helps to settle the debate. Although there are cases where government intervention coexists with success, there are many instances where industrial policy has failed to yield any gains. The most difficult issue is that relevant counterfactuals are not available. Consider the argument that Japan's industrial policy was crucial for its success. Because we do not know how Japan would have fared under laissez-faire policies, it is difficult to attribute its success to its industrial policy. It might have done still better in the absence of industrial policy — or much worse. Given this basic difficulty, only indirect evidence can be obtained regarding the efficacy of industrial policy. Direct evidence that can "hold constant" all the required variables (as would be done in a well-specified econometric exercise) does not exist and likely never will.

The following section analyzes the main conceptual arguments in favor of industrial policy. Because the infant industry argument for trade protection anticipates most of the rationales for industrial policy, this argument is discussed extensively. Next, India's successful software industry is examined in the context of industrial policy. The following section examines how the expansion of international production networks has altered the case for industrial policy. The final section considers the issue of policy space.

## Why Industrial Policy?

At a general level there is room for government intervention when there are market distortions (such as externalities or market power) or when markets are incomplete (for example, future markets for many goods simply do not exist). As is known from a basic theorem of welfare economics, under such market failures a competitive market system does not yield the socially efficient outcome. Any argument for industrial policy is a special case of this general argument.

Three specific arguments for industrial policy have received the most attention. One is derived from the presence of knowledge spillovers and dynamic scale economies, a second from the presence of coordination failures and a third from informational externalities. Before discussing these arguments in detail, it is useful to begin with the infant industry argument for trade protection because it is a precursor of modern arguments for industrial policy.

## *The Infant Industry Argument: A Precursor of Modern Industrial Policy*

The infant industry argument is one of the oldest arguments for trade protection and perhaps the only such argument that is not dismissed out of hand by economists. The most popular (and the simplest) version of the argument runs as follows. Production costs may initially be higher for newly established domestic industries than for well-established foreign competitors, which have more experience. Over time, however, domestic producers can reduce costs as they learn by doing (they enjoy dynamic scale economies), and they can eventually attain the production efficiency of their foreign rivals. However, if the fledgling domestic industry is not initially protected from foreign competition, it may never takeoff. Furthermore, if dynamic scale economies are strong enough, temporary protection of the domestic industry can be in the national interest.

A stronger version of the argument states that the domestic industry might even be capable of attaining production costs below its foreign rivals if it is given sufficient protection. In this version of the argument, true comparative advantage lies with the domestic industry, and temporary protection can be in the global interest, because consumers in the rest of the world also benefit from the eventual lower production cost of the domestic industry.

In an influential paper, Baldwin (1969, p. 297) provided an incisive criticism of the infant industry argument, contending that "if after the learning period, unit costs in an industry are sufficiently lower than those during its early production stages to yield a discounted surplus of revenues over costs (and therefore indicate a comparative advantage for the country in the particular line), it would be possible for firms in the industry to raise sufficient funds in the capital market to cover their initial excess of outlays over receipts." If future returns indeed outweigh initial losses, capital markets would finance the necessary investment needed by the domestic industry. It is obvious, but worth stressing, that if future returns fall short of initial losses, the industry should not be established in the first place.

A frequently cited counter to Baldwin (and one that he acknowledged) is that capital market imperfections might prevent the infant industry from obtaining the required financing. For example, because of informational asymmetries investors, unlike producers, may not know that the industry will be profitable in the long run and therefore fail to provide the capital needed to cover the initial costs. However, such an argument defies credibility because it requires one to believe that firms that have not even begun production know more about their prospects than investors whose main objective is to find profitable uses for their excess capital and have previously analyzed and financed similar projects. And even if one grants the presence of asymmetric information, what prevents potential producers from conveying such information to likely investors? After all, entrepreneurs seeking funds for new businesses overcome this problem on a routine basis.

While the infant industry argument assumes that it is known with certainty that the industry in question will eventually be profitable, it seems more likely that the prospects for most new industries are uncertain and that *no one* really knows whether a particular infant industry will be profitable in the future. Under such circumstances capital markets would require compensation for the risks involved, and the resultant interest rates required might make the investment unprofitable. But efficiency requires that those bearing risks should be compensated, and there is no market failure if the underlying problem is that investors do not provide the necessary capital because they perceive the rewards not to be commensurate with the risks they are asked to bear.

Nevertheless, the assumption of omniscient financial intermediaries should be viewed with some skepticism. From early bubbles such as the tulip mania of the seventeenth century to the Internet bubble of the late 1990s, it is clear that financial actors are often deficient. In the Asian countries that suppressed the financial sector and directed loans to specific industries and firms as a part of industrial policy, the banking sector was itself in need of substantial improvement in operating procedures, much as industrial firms were. Thus, the belief that if there were opportunities investors would exploit them might be somewhat of a weak link in Baldwin's argument. On the contrary, it also implies that any selective economic policies would have to simultaneously address the weakness of the financial sector along with that of goods and other services. Indeed there might be an argument for initially strengthening the banking sector, perhaps by allowing foreign financial intermediaries into the country, before pursuing targeted sectoral policies. In any case, as Baldwin notes, if there is a problem with capital markets, policy ought to narrowly target that problem rather than resort to trade protection.

In today's world of global capital markets the simple version of the infant industry argument runs into another difficulty: Investors ought to be able to determine the prospects for the domestic infant industry from the experience of foreign producers. If domestic investors lack such information, surely foreign investors ought to have it. Why cannot the borrowing be international rather than local? One answer to this question may be that investors believe that an industry that has succeeded abroad will not necessarily succeed at home. But this explanation can be consistent with the very hypotheses underlying the infant industry argument only if investors are not fully rational.

What light has formal analysis shed on the infant industry argument? A seminal paper by Bardhan (1971, p. 1) noted that the infant industry argument is dynamic and that "any elaboration of this idea involves explicitly dynamic analysis, and it has hardly been integrated into the main corpus of trade theory which is mostly comparative-static in nature." Bardhan provides the first dynamic model of learning by doing in an open economy and derives the optimum extent and time path of protection to the learning industry. His model has two goods,  $c$  and  $m$ , and two factors of

production, capital and labor, with constant returns to scale in production of both goods. The learning effect is assumed to depend on the cumulated volume of industry output in good  $m$ , and it shifts out the production function for the good in a Hicks neutral fashion.'

Bardhan models learning by doing as a classic Marshallian externality: The higher the cumulative output of the industry, the more productive is the technology of each individual firm. When learning is unbounded, Bardhan shows that it is socially optimal to subsidize the infant industry and that the time profile of the optimal subsidy depends on initial conditions. However, his framework does not capture the idea that international spillovers may partially substitute for domestic learning because the learning effect function contains the stock of domestic and foreign outputs as separate arguments, and the relationship between the two is not really considered.<sup>2</sup>

Succar (1987) extends Bardhan (1971) analysis to allow the learning in one sector to generate spillovers for both sectors, thereby providing an interindustry spillover rationale for the infant industry argument. However, the presence of such economies is not sufficient to justify intervention. As Succar notes, the discounted stream of productivity gains generated by learning by doing in the infant industry should outweigh the discounted stream of subsidies or else intervention is socially undesirable.<sup>3</sup> The intuitive idea underlying Succar's model is that the production of capital goods can enhance growth by acting as an "informal learning center where technical skills are required" thereby contributing to a country's technical infrastructure (p. 523).<sup>4</sup> Such improvements in the skill base of workers complement investments in human capital and can advance industrialization in developing countries.

The distinction between firm- and industry-level learning by doing becomes quite important because firms are heterogeneous. Suppose that some firms are more efficient at learning than others. Optimal subsidies would have to be nonuniform, and the government is unlikely to possess the information needed to implement an optimal subsidy program. It might thus make sense for the government to adopt a uniform policy even though that might not be the first-best policy. While in theory, mechanisms could be designed that induce firms to reveal their learning capabilities, the practical relevance of such mechanisms is far from clear.

As might be expected, there is more to the infant industry argument than the simple version formalized by Bardhan and Succar. As Baldwin notes, there are four more nuanced versions of the infant industry argument. First, acquisition of knowledge involves costs, but knowledge may not be appropriable by an individual firm. This is the standard argument for subsidizing research and development (R&D). Second, firms may provide costly on-the-job training but may be unable to prevent the diffusion of such knowledge as workers move to other companies (a free-rider problem in worker training). While firm-specific training involves no potential externality,

general training can lead to externalities that would justify subsidies. Third, static positive externalities in the production of a good may justify trade protection. And fourth, determining the profitability of a new industry might require a costly investment, and the results could become freely available to potential competitors. In other words, investment in new industries might result in informational externalities that make it difficult for investors to earn a rate of return high enough to justify the initial investment. This argument has been formalized by Hausmann and Rodrik (2003), who call it the process of self-discovery, of determining what a company can produce profitably at world prices.

The infant industry argument does not really specify how learning occurs. It just assumes that dynamic scale economies will somehow be realized by the infant industry. Of course, learning is rarely exogenous, and it usually requires considerable effort and investment by firms (Pack and Westphal 1986). If such investments are made, firms need to be able to appropriate the benefits of the knowledge gained. Knowledge is a nonrival good and, once created, any number of agents can use it simultaneously. If firms cannot prevent the leakage of knowledge that is costly to create, they will have little incentive to create such knowledge. If property rights over knowledge are not enforceable, this can create a rationale for government intervention.

As Baldwin (1969) notes, many types of knowledge acquisition are not subject to the externality described above because entrepreneurs can often prevent the leakage of their knowledge to potential competitors. Similarly, if there are only a few firms in the industry, interfirm negotiations should help offset the externality problem (Coase 1960). But what if many rival firms benefit from the investment undertaken by a knowledge acquiring firm, and the firm can do nothing to prevent such diffusion? Is government intervention justified?

Trade protection is certainly not called for. A tariff does nothing to solve the basic externality problem and may well worsen it. A production subsidy to the entire sector will also fail to remedy the externality. What is needed are subsidies to initial entrants into the industry that help create new knowledge and discover better production technologies. As with R&D subsidies governments should target the marginal rather than inframarginal research. In the case of new firms it takes time to discover whether a new idea or technology is socially valuable, and the adoption of a novel technology by others is the strongest proof of its social value. Thus, a policy of rewarding early entrants requires an accurate forecast of the social value of their inventions and discoveries—a process that can be fraught with failure. Not only that, given the uncertainty associated with new technologies, a delayed pattern of adoption might even be socially optimal.

### *Knowledge Spillovers, Dynamic Scale Economies, and Industrial Targeting*

Ever since David Ricardo, it has been well known that under free trade a country can increase its national income (and welfare) by moving resources into sectors in which

its opportunity cost of production is lower than that of its trading partners. But is this prescription sufficient to generate economic growth? Perhaps not. Allocating resources according to comparative advantage can only ensure static efficiency and in no way guarantees dynamic efficiency. Sussar (1987, pp. 533–34) argues that "the comparative advantage theory is a static construct that ignores [that] forward linkages exist between present choices and future production possibilities. Therefore it cannot guide the pattern of international specialization when there are asymmetric learning opportunities associated with the production of different goods, use of certain techniques, or both. Promotion of industries which generate substantial learning by doing economies should be an integral part of a strategy of human capital formation in [developing countries]." In other words, Sussar argues for some sort of industrial targeting, although her model does not explicitly deal with this issue.

Even if one accepts the premise that certain industries are more likely to generate spillovers (based on knowledge diffusion or other factors), can policy be designed to encourage the "right" industries? The ideal but rarely attained goal of industrial policy is the development of a general-purpose technology. The Defense Advanced Research Projects Agency (DARPA), a small unit within the U.S. Department of Defense that generated and financed a portfolio of projects, is widely credited with having been the key contributor to the development of the Internet, in response to the need to maintain communications during an assault on the United States. This instance of success addressed a market failure, in that the social benefits of the research were much larger than the anticipated private benefits. Moreover, DARPA foresaw a potential need that private firms might not have. While the Internet was a main technological breakthrough and suggests the potential gains from such activity, it is useful to remember that the discovery of such general purpose technologies is a rare event and less likely in low innovation-intensity developing countries than in research-rich industrial countries.

The informational constraints facing policymakers pursuing industrial policy are severe and any realistic model of industrial targeting needs to account for them. In a recent paper Klimenko (2004) models industrial targeting as an optimal experimentation strategy for a government that lacks information about the set of industries in which the economy has a comparative advantage. He examines the set of industries in which a country will specialize as a result of such policy. In his model, for any set of targeted industries, it is possible to know with positive or zero probability whether a country will specialize in this set. He shows that an optimally designed industrial policy can lead a country to specialize in sectors in which it does not have comparative advantage. Depending on the beliefs of the policymaker, a country can end up abandoning the industries in which it has "true" comparative advantage.

Furthermore, Klimenko argues that policymakers may stop looking for better targets when the favored industries perform well enough. He interprets this outcome as a failure of industrial targeting policy even though it may not appear to be. He

goes on to show that despite the existence of market failures, the outcome of the learning process through private experimentation (without any assistance from the government) can yield outcomes that are closer to the full information social optimum. Klimenko's rigorous analysis underscores the intuitive argument that the relevant counterfactuals are unavailable, and what may appear to be a successful industrial policy may not be the first-best outcome from a country's perspective. Merely doing something well need not imply that one might not be better at something else.

### *Coordination Failures as a Rationale for Industrial Policy*

The idea behind the coordination failure argument for industrial policy is that many projects require simultaneous investments to be viable, and if these investments are made by independent agents, there is little guarantee that each agent, acting in its own self-interest, would choose to invest.<sup>5</sup> As Scitovsky (1954) noted, reciprocal pecuniary externalities in the presence of increasing returns can lead to market failure, because the coordination of investment decisions requires a signaling device to transmit information about present plans and future conditions, and the pricing system is not capable of playing this role.

Paoli and Westphal (1986) argue that such pecuniary externalities related to investments in technology are pervasive during industrialization. They provide an example of two infant industries, where industry A produces an intermediate that is required in industry B and neither industry is profitable if it is established alone. However, if both industries are established together, both are profitable, implying that it is socially optimal to establish both. Of course, the problem is that without explicit coordination of investment decisions, this outcome would not be obtained.

Okuno-Fujiwara (1988) presents a formal model of such interdependence between industries and the coordination failure that can result. He considers an economy with three goods,  $x$ ,  $y$ , and  $z$ , where good  $z$  serves as a numeraire and is produced under perfect competition with constant returns to scale. Good  $x$  is produced by a competitive industry and requires good  $y$  as an intermediate. The technology for good  $y$  exhibits large economies of scale, and the industry is assumed to be oligopolistic with the number of firms endogenously determined to ensure zero profits in equilibrium.

A coordination problem arises in the industry because the derived demand for the intermediate good  $y$  depends on its price, which in turn determines incentives for entry into the intermediate sector. If  $y$  producers anticipate low demand for their good, given the fixed costs of entry, few new producers would want to enter, implying a higher price for the intermediate. This could make industry  $x$  unsustainable. The key assumption here is that the intermediate good  $y$  must be locally supplied. On the contrary, if  $y$  producers are certain of high demand for their product, more would

enter, lowering its price and allowing the high demand for the intermediate to be sustained.<sup>6</sup> Okuno-Fujiwara (1988) shows that there is no unique equilibrium in a small open economy with the above production structure. In the bad equilibrium the economy ends up specializing in good  $z$  whereas in the good equilibrium it produces both goods  $x$  and  $y$  and exports good  $x$  to the rest of the world (where the good equilibrium is welfare superior to the bad).

Turning to policy analysis, Okuno-Fujiwara (1988) suggests that three types of traditional government intervention can help ensure that the good equilibrium is realized: The government can provide a production subsidy to industry  $x$  or industry  $y$  or both, causing the two sectors to expand; it can provide an export subsidy to industry  $x$ ; or it can shutoff international trade. However, he notes that trade protection can be effective only if the autarkic equilibrium production of good  $x$  is sufficiently large—something that is less likely to be true of small developing countries. In addition to traditional industrial policies (the first two options), Okuno-Fujiwara (1988) also suggests that the government can play a coordinating role between  $x$  and  $y$  producers by facilitating information exchange.<sup>7</sup> However, he argues that only repeated information exchanges can resolve the coordination failure. It is difficult to believe that policymakers can effectively execute such information exchanges between industries about whose day-to-day business they likely know little. Furthermore, this policy prescription suggests a massive role for government intervention in industrialization. Okuno-Fujiwara himself is skeptical of whether the mechanisms captured by his model and the policy prescriptions that emerge from his analysis had any practical analog in the Japanese experience.

In a paper along the lines of Okuno-Fujiwara (1988), Rodrik (1996) argues that for coordination failures to exist between upstream and downstream industries there must be some type of scale economies in production and imperfect tradability must hold across national borders for some of the goods, services, or technologies associated with manufacturing. In his model the intermediate good sector is characterized by monopolistic competition rather than oligopoly. He suggests that the non-tradable intermediate goods sector should be viewed as representing different categories of specialized skill labor. The idea is that a worker's decision to acquire a skill depends on demand for that skill and that it is costly or simply infeasible to import labor services should certain skills be in short supply locally. Like Okuno-Fujiwara (1988), Rodrik (1996) is hesitant to offer strong policy recommendations based on his analysis and concludes that government intervention designed to resolve such coordination failures is a risky strategy. The World Bank's (1993) well-known report on the East Asian miracle argues that East Asian efforts to coordinate investment decisions led to a number of inefficient industries.

While the theoretical rationale for redressing coordination failure appears to be sound, the argument rests on certain key assumptions, particularly that the organization of production activity is exogenously given. Why would industries whose

profitability is so intimately intertwined not find ways to help coordinate decisions, as is the case in the many international supply networks (Sturgeon and Lester 2002, 2003; Gereffi and Memedovic 2003)? For example, vertical integration between intermediate and final goods producers can help resolve some coordination problems, although there are clearly limits to the extent to which organizations can adjust their scale and scope to solve coordination problems. At some point all firms have to interact with others in the market. But long-term contracts between firms have been used to solve problems of relation-specific investments in many industries. It is not clear why contracts could not play the same role for coordination failures.

Perhaps the biggest problem with the coordination failure argument is that it relies heavily on the assumption of nontradable intermediate inputs, partly reflecting the fact that much of the early literature was based on the example of the steel and automobile industries of the 1960s rather than products for which transportation costs for intermediates are likely to be low. Virtually all the models make this assumption although most international trade is in intermediate goods. Thus, the coordination failure argument runs up against the central fact about which much of the "new" trade theory has been built (see, for example, Ethier 1982).

This is no small contradiction, and if the coordination failure story is to be rescued, it needs to appeal to nontradable services as in Rodriguez-Clare (1996b). But the problem then is that the case for industrial policy on the basis of coordination failures is quite thin if inward foreign direct investment (FDI) is feasible. If local firms do not produce sufficient intermediates because of coordination failures, why could intermediates not be produced by foreign multinationals that are surely not dependent on the production structure of any one economy? In small developing countries a large-scale investment by a multinational can create sufficient demand for intermediates and easily resolve the coordination problem. This is partly what the literature on the backward linkage effects of FDI argues (Rodriguez-Clare 1996a; Marlcusen and Venables 1999).

It is unlikely that multinational firms would experience the type of coordination problems that confront small producers in developing countries. Indeed, the huge growth in international supply chains established by multinational firms has become one of the most visible features of industrial growth in the last decade (Sturgeon and Lester 2002). The role of multinational firms in determining the overall case for industrial policy is discussed in more detail below.

### *Informational Externalities*

In a recent paper Rodrilc (2004) argues that the traditional view of industrial policy (based on technological and pecuniary externalities) does not capture the complexities of industrialization. He argues that industrial policy is more about eliciting

information from the private sector than it is about addressing distortions through first-best instruments. He envisions industrial policy as a strategic collaboration between the private and the public sectors with the primary goal of determining the areas in which a country has comparative advantage. The fundamental departure of this viewpoint from classical trade theory is that entrepreneurs may lack information about where a country's comparative advantage lies. Or more to the point, at the microlevel, entrepreneurs may simply not know what is profitable and what is not.

In the presence of informational externalities, a free-rider problem arises between initial and subsequent investors. Suppose no one knows whether activity  $x$  is profitable and that the uncertainty can be resolved only by making a sunk investment. By definition the investment cannot be recovered if the outcome turns out to be unfavorable. If there is free entry *ex post*, no entrepreneur may be willing to make the required investment *ex ante*: If the activity indeed turns out to be profitable, other entrepreneurs will be attracted to it, thereby eliminating all rents.

It is worth noting that Baldwin's (1969, p. 302) classic paper anticipates Rodrik's argument almost exactly: "suppose, for example, that a potential entrant into a new industry, if he could provide potential investors with a detailed market analysis of the industry, could borrow funds from investors at a rate that would make the project socially profitable. However, should this information become freely available to other investors and potential competitors, the initial firm might not be able to recoup the cost of making the market study....under these circumstances the firm will not finance the cost of the study, and a socially beneficial industry will not be established." Similarly in the context of the adoption of high yielding varieties of crops by farmers in developing countries, Besley and Case (1993, p. 399) note that late adopters may learn from early adopters: "[when] a technology is of uncertain profitability, some potential adopters may wait until they observe whether others have fared well by using it" and that such "externalities are potentially important in agricultural technology adoption."

Given the importance of this argument for the debate on industrial policy, it is useful to examine the framework presented in Hausmann and Rodrik (2003) in some detail. They consider a small open economy with two sectors, traditional and modern. The production technology in the traditional sector is constant returns to scale, and the presence of a fixed factor generates diminishing returns. In the modern sector, which consists of many goods, there are constant returns to scale in production, but the cost of production of a good depends on an unobserved productivity parameter,  $\theta_i$ , that becomes known only when the production of a good is attempted. This requires a time period in which resources must be used but no production takes place—what Baldwin called a "market study." Entrepreneurs lack information about the profitability of production of various goods in the modern sector, and this information can be obtained only by undertaking a sunk investment.

After uncertainty regarding  $\theta_i$  is resolved, entrepreneurs compare their production costs with world prices and produce goods for which they make monopoly profits, which accrue for length of time  $T$ —call this the monopolization period. Of course, once information becomes public, which it does in period three when the monopolization period has elapsed, there is further entry into goods that yield positive profits until all profits are competed away to zero.

Hausmann and Rodrik (2003) analyze the laissez-faire equilibrium of the above model and compare it with the social planner's problem to derive the market failures that result from the presence of informational externalities. They argue that the market equilibrium is deficient in two respects. First, the level of investment and entrepreneurship delivered by the market does not coincide with the social optimum, because the entrepreneurs care only about profits and not about the economy-wide benefits of their investment. If the monopolization period is long, the market economy can actually deliver too much investment in the modern sector rather than too little. This suggests that in economies where firms face substantial entry barriers, the underinvestment problem noted by Hausmann and Rodrik (2003) is not likely to be serious. For example, the industrial licensing regime pursued by India during the first 40 or so years after independence made it difficult for firms to enter new markets. And the recent literature on the business climate emphasizes other factors that discourage investment in the modern sector, such as the time to obtain business permits, telephone lines, and other utility hookups (World Bank 2006). Such barriers should have helped protect rents for those that did manage to enter profitable markets.

The second market failure identified by Hausmann and Rodrik (2003) is that the market equilibrium yields too little specialization—all activities that turn out to be profitable are sustained whereas optimality requires that only the one activity with the highest return be pursued. In other words, while it is optimal in their model for the small open economy to produce only the good for which the profit margin is the highest, the market solution allows all those that make positive profits to stay in business during monopolization period.

This result reflects the general equilibrium nature of their model and the fact that they consider a small open economy. To see this, first note that the modern sector draws resources out of the traditional sector and that optimality requires that these resources be used where they generate the largest profits, which happens to be in the modern good for which the productivity parameter ( $\theta_i$ ) is the highest. Second, because the country's output of a good does not affect the world price, one can never have a situation where the markups across different goods are equalized. Clearly, if world prices changed with a country's exports or output, complete concentration in the modern sector need not obtain. A more likely scenario would be that a country should produce higher quantities of modern goods for which it has a more favorable productivity draw and lower quantities of other goods.

Hoff (1997) argues that if initial producers benefit subsequent producers, the case for subsidizing initial producers hinges much on the assumption that the externalities operate in a deterministic fashion (do not involve any uncertainty). She constructs a model in which initial entrants provide information that is socially valuable by reducing uncertainty for potential followers regarding production conditions. In her model factors that increase the informational barrier to entry can actually imply a lower optimal subsidy for the infant industry. By contrast in most models the externalities are assumed to remove all uncertainty rather than simply reducing it. Because Hoff's model is clearly more realistic, it is notable that her results weaken the case for subsidizing an infant industry.

### *The International Dimension: Role of Exports and FDI*

For small developing countries the case for industrial policy is rarely a purely domestic one. International considerations are fundamental, and the role of exports (on the part of domestic firms) and inward FDI has received considerable attention. A potential rationale for industrial policy in the context of exports arises when product quality is unknown to foreign consumers. The information asymmetry can lead to market failure that can then potentially justify some form of intervention. Adding an explicit process of reputation acquisition may be an objective of policy. Grossman and Horn (1988) focus on reputation acquisition at the firm level, whereas Mayer (1984) focuses on the country level. In the view of Grossman and Horn, Toyota can affect only its own reputation in foreign markets whereas in Mayer's model, experience with Toyota also influences how foreign consumers view other Japanese companies, such as Honda. The difference matters because returns to reputation acquisition are appropriable in the Grossman and Horn model whereas they are not in the Mayer model.'

Policy intervention with respect to FDI has a long history. The rationale has frequently been the effects of FDI on the productivity of local firms through technology transfer and linkage effects. The literature on FDI, technology transfer, and linkages is extensively surveyed by Saggi (2002). The review here is limited to aspects of FDI that relate intimately to local industrial development and its linkage effects, because these correspond quite well to the coordination failure rationale for industrial policy.

There is a voluminous informal as well as empirical literature on backward linkages. For example, the *World Investment Report 2001* (UNCTAD 2001) was devoted entirely to the effects of FDI on backward linkages in host countries. However, analytical models that explore the relationship between multinationals and backward linkages in the host country are hard to come by. Two examples of such models are Markusen and Venables (1999) and Rodriguez-Clare (1996a). Both models emphasize the demand-creating effects of FDI on the host economy: Multinationals generate derived demand for intermediate goods, thereby promoting industrial development

of the intermediate goods sector in the host country.<sup>9</sup> As noted, a common problem with analytical models in this area is the assumption that intermediates are nontradable. These models assume no trade in intermediates and then use FDI as the channel that provides some intermediates or increases demand for local intermediate goods producers. As a result, the models are likely to overstate the impact of multinationals on industrial development.

Mexico's experience in the automobile industry is illustrative of how FDI can contribute to industrial development in the host country (Laderman, Maloney, and Serven 2003). Initial investments by U.S. car manufacturers into Mexico were followed by investments by Japanese and European car manufacturers and automobile parts and component manufacturers. As a result, competition in the automobile industry increased at multiple stages of production, efficiency improved, and Mexican automobile industry exports boomed. The pattern of FDI behavior in Mexico—investment by one firm followed by investment by others—probably reflects strategic considerations involved in FDI decisions. Most multinational firms compete in concentrated markets and are highly responsive to each other's decisions. An important implication of this interdependence among competing multinationals is that a host country may be able to unleash a sequence of investments by successfully inducing FDI from one or two important firms. However, the concentration of inward FDI into a handful of developing countries suggests that only a few countries can benefit from this process—Egypt and Tanzania are not China.

A recent case study of the effects of Intel's investment in Costa Rica by Larrain, Lopez-Calva, and Rodriguez-Clare (2000) finds evidence that local suppliers benefited substantially from Intel's investment. Similar evidence exists for other sectors and countries and is discussed in Moran (1998, 2001). For example, in the electronics sector in Malaysia, Moran (2001) notes that foreign investors helped their local subcontractors keep pace with modern technologies by assigning technicians to the suppliers' plants to help set up and supervise large volume automated production and testing procedures. In a broader study Batra and Tan (2002) use data from Malaysia's manufacturing sector to study the effect of multinationals on interfirm linkages and productivity growth during 1985–95. Their results show that not only are foreign firms more involved in interfirm linkages than domestic firms but also that such linkages are associated with technology transfer to local suppliers. Such technology transfers were found to have occurred through worker training and the transmission of knowledge that helped local suppliers improve the quality and timeliness of supply.

Javorcik (2004) examines backward linkages and technology spillovers using data from the Lithuanian manufacturing sector during 1996–2000. She finds that firm productivity is positively affected by a sector's intensity of contacts with multinational customers but not by the presence of multinationals in the same industry. Thus, her results support vertical spillovers from FDI but not horizontal spillovers.

Furthermore, she finds that vertical spillovers occur only when the technological gap between domestic and foreign firms is moderate. Blalocli (2001) uses a panel data set from Indonesian manufacturing establishments to check for the same effects. He finds strong evidence of a positive impact of FDI on productivity growth of local suppliers, showing that technology transfer does take place from multinationals. He also plausibly suggests that because multinationals tend to source inputs that require relatively simple technologies relative to the final products they produce, local firms that manufacture such intermediates may be in a better position to learn from multinationals than those that compete with them.

If one accepts the optimistic view of the effects of FDI—and some of the evidence discussed above suggests reasonable grounds for doing so—does this have implications for industrial policy? The answer is a qualified yes. Basic economic theory tells us that it is optimal to subsidize an activity if it generates positive externalities—if the activity benefits agents other than those directly involved in the activity. The potential surely exists for positive externalities from FDI, and evidence exists that this potential is often realized. Incentives to attract FDI may be justified on the grounds of such externalities from inward FDI, but the magnitude of some of the incentives being used seems difficult to justify (Moran 1998), and such policies are not typically what proponents of industrial policy have in mind. Indeed, the thrust of such arguments is typically in favor of encouraging the development of indigenous firms. Investment incentives and tax breaks to multinational investors work against their local competitors. Thus, if there are local firms that could potentially compete with multinationals, the adverse effect on such firms of tax incentives to multinationals needs to be taken into account. The efficacy of investment incentives is also unclear—such policies could easily end up transferring rents to foreign investors without affecting their investment decisions.

### *Government Knowledge Requirements*

This review of arguments for industrial policy suggests the enormous difficulties of implementing industrial policies quite apart from the possibilities for rent-seeking. The range and depth of knowledge that policymakers would have to master to implement successful policy is extraordinary. They would have to be accurately informed about an enormous range of complex questions, understand their relevance, and be able to accurately evaluate subtle differences. Some of the issues on which policymakers would have to be knowledgeable derived from the preceding discussion include:

- The firms and industries that generate knowledge spillovers.
- The firms and industries that benefit from dynamic scale economies—the precise path of such learning and the magnitude of the cost disadvantage at each stage of the learning process.

The sectors that have a long-term comparative advantage.

The size of scale economies of different firms and sectors, to facilitate investment coordination.

- An ability superior to that of individual firms to learn about their potential competitiveness.
- The nature and extent of capital market failures.
- The magnitude and direction of interindustry spillovers.
- The relative amount of learning by individual firms from others and from their own experience.
- The extent to which early entrants generate benefits for future entrants.
- The extent of heterogeneity of firms' learning abilities.
- Whether consumers learn the quality of a good after consuming rather than by inspecting it.
- Whether firms that are trying to reduce production costs also begin a simultaneous effort to improve their product's quality to obtain a better reputation.
- The potential effects of FDI or international trade on coordination problems, including a detailed knowledge of which of tens of thousands of intermediates are tradable.
- A forecast of which firms can create new knowledge and discover better production methods.

The spillover effects of FDI and the likely intensity of foreign purchase of domestic intermediates.

It is possible that government officials might be this omniscient, but the performance of the portfolio managers in industrial country stock markets suggests that few of the well-trained (and remunerated) equity analysts can evaluate even much more certain and grosser characteristics of existing firms and industries with long track records. Nor do industrial firms themselves have the ability to successfully forecast such developments. Acknowledging that a first-best policy would argue for the government to address such market failures or externalities, the task is daunting. Quite apart from the dangers of optimal policy being subverted by industries and firms that would benefit, the sheer knowledge and skill requirements would exceed that possessed by almost any institution, including the best consulting firms. On a far more circumscribed set of tasks, measuring and explaining the sources of lower total factor productivity for a small number of sectors in Brazil and the Republic of Korea relative to the United States, McKinsey & Co., a preeminent consulting firm, spent several years and employed dozens of people with qualifications exceeding those of officials in most developing countries (McKinsey Global Institute 1998a,b).

No study has attempted to assess whether governments have mastered these 15 areas (or others that can be derived from the discussion here) that have to be addressed. The efficacy of industrial policy has to be evaluated on the basis of the

realized results of the firms or industries that have been encouraged. The underlying market failures or externalities that contributed to the decision to foster a firm or sector cannot be identified from the policy (such as subsidized directed credit). Only the effects of the policy can be assessed. This task is taken up next.

## Does Industrial Policy Work?

As noted, it is impossible to offer a single agreed counterfactual to evaluate the success of industrial policy targeted to individual industries. Thus there have been a number of research strategies to provide an empirical evaluation of industrial policy. These are reviewed in Noland and Pack (2003). Among other issues, researchers have examined the impact of trade protection, subsidies to R&D, general subsidies, and preferential lending rates on the evolution of productivity, capital accumulation, and sectoral structure. Few of the empirical analyses find that sectoral targeting has been particularly effective.

Consider some of the evidence. In Japan more than 80 percent of on-line budget subsidies were devoted to agriculture, forestry, and fisheries in 1955–80, the peak of Japan's industrial policy efforts.<sup>10</sup> Implicit tax subsidies for investment were highest in the mining sector and low in the high technology sectors. Government subsidies to R&D were also small. Unless elasticities of investment and R&D with respect to subsidies were implausibly high, their effect was limited. Industries that were encouraged did not experience significantly faster rates of total factor productivity growth than others, and R&D subsidies were largely ineffective.

Beason and Weinstein (1996) examine the connection between industrial policy and sectoral total factor productivity growth in Japan. Working with a 13-sector sample for 1955–90, they fail to uncover evidence that preferential policies (measured by the effective rates of protection, taxes, or subsidies) targeted sectors with increasing returns to scale or contributed to the rate of capital accumulation in targeted sectors or to their total factor productivity growth. They do find some evidence that before the first oil shock, industrial policy targeted sectors with high labor usage. Employing a slightly different data set, Lawrence and Weinstein (2001) extend this research and find that differential corporate tax rates had an impact on sectoral total factor productivity growth, whereas direct subsidies and subsidized loans did not. Moreover, they find the paradoxical result that the effective rate of protection was negatively associated with sectoral total factor productivity growth and that imports, not exports, were positively associated with total factor productivity growth.

There are at least two channels through which imports could contribute to increasing productivity. First, imports allow domestic producers to use new, improved, or specialized intermediate inputs to which they would not otherwise

have access. Second, imports compete with domestic products, and their availability acts as a constant spur to domestic producers to cut costs and improve quality. Lawrence and Weinstein (2001) divide imports into "competitive" and "noncompetitive" and find evidence for Japan to support this second channel. From this they conclude that Japan's growth would have been even faster if it had cut tariffs and exposed a greater share of its domestic producers to foreign competition. \*\*

Following a method broadly similar to that of Beason and Weinstein (1996), Lee (1996) finds a similar lack of impact of Korean industrial policies on sectoral capital accumulation or total factor productivity growth. Pack (2000) follows a different strategy, assuming that total factor productivity increased in favored manufacturing sectors in both Japan and the Republic of Korea and estimates how much of an impact even an assumed successful policy could have had on the growth of gross domestic product. The most favorable estimate is a roughly 0.5 percentage point increase in total gross domestic product growth rate of roughly 10 percent over the relevant periods. While this is significant, it is hardly the magic key to accelerated growth.

It is possible that the impact of industrial policy is manifest largely in sectors that purchased inputs from the promoted sectors, even if the promoted sectors did not themselves benefit. However, Pack (2000) finds that sectors that were encouraged in Japan and the Republic of Korea had few linkages with nonfavored sectors through input-output relations, and there is little evidence of labor flowing from favored to neglected sectors, a likely mechanism for the transmission of knowledge.

Nevertheless, as noted at the beginning of this article, the difficulty of constructing a single agreed on counterfactual precludes a robust conclusion. Moreover, **all** the empirical analysis examines the contemporaneous impact of policies—for example, did Korean industries that were encouraged experience greater total factor productivity growth in the period when main promotion occurred, 1973–85? Someone doubting these results could point to the performance of Korean firms such as Samsung and LG in the following two decades in such diverse product lines as plasma televisions, RAM chips, and cellular phones and attribute these later successes to the earlier stimulation the firms received for other product lines. These more recent successful efforts by the firms could be attributed, in this interpretation, to their earlier growth in other product categories. In this view learning to perform R&D on microwaves had future carryover effects on plasma televisions. Fully resolving divergent views is impossible, but detailed firm histories by Kim (1997) or Hobday (1995) do not suggest such carryover.

Even if it could be shown that the success of a few firms is attributable to earlier encouragement by the government, the aggregate effects cited above suggest there was not a major impact at the national level during the main period of growth acceleration. And any such effects would have to be weighed against the negative long-run impacts in the financial sector cited by those skeptical of industrial policy. For

example, the Asian financial crises of the late 1990s and Japan's stagnation since 1990 can be interpreted as partly the result of earlier government directed lending that minimized the need for banks to learn modern techniques of evaluating individual projects and managing the riskiness of their overall portfolio.

## New Industrial Policy

Recent discussion of "new" industrial policy including the desirability of fostering learning and obtaining benefits from agglomeration economies offered by industrial clusters has received little systematic empirical evaluation. Rodriguez-Clare (2004a,b) provides an extensive discussion and a formal treatment of clusters. Humphrey and Schmitz (2002) provide an extensive survey of the empirical literature on clusters and discuss whether they offer a locally controlled alternative to participation in networks.

### *Export Processing Zones and Other Clusters*

In principle, the development of clusters could boost productivity through the provision of overhead services by the organizers plus the interaction of the firms entering the cluster. Clusters could offer an alternative to dependence on either buyer- or manufacturer-led networks.

The benefit of clusters may arise from face-to-face interactions that are productivity enhancing (interactions between software writers and chip manufacturers, for example), a pool of workers with the relevant skills, or reduced transportation costs. Individual market agents may not be aware of the externality they generate for others, and this provides an additional market failure that could in principle be addressed by public intervention. The main example usually cited is that of Silicon Valley in California, which most accounts suggest arose spontaneously. Similarly, the rapid development of the software industry in Bangalore and other cities in India, discussed below, appears to be the outcome of the combination of a large group of well-educated English-speaking students, the entrepreneurial abilities of a small group of residents, and the activities of a large Indian expatriate community, particularly in Silicon Valley. While publicly financed education institutions generated the fundamental resource, educated workers, there was no explicit effort to galvanize the agglomeration economies that developed. Texas Instruments financed a critical communications satellite. Positive government efforts followed the takeoff of the sector.

There are interesting descriptions of a number of clusters in high-income countries, but few normative evaluations of their success employing social cost-benefit analyses or even grosser measures such as growth of exports relative to firms outside

the cluster but in the same sector. However, some insights can be obtained about whether recent success stories in Asia conform to the contours of the new industrial policy.

Development of the Indian software sector reflected a complex set of interactions between domestic and foreign responses to perceived opportunities. The evolution of the Indian software industry centered in Bangalore is explained in detail in the following section. Many of the same patterns, with different details, can be documented for other success stories, such as the Hsinchu Science Park in Taiwan, China (Saxenian 1999, 2001), the special economic zones in China (Rosen 1999; Huang 2003), and Bangladesh's rise as a clothing exporter (Rhee 1990). In the Indian software sector and the Bangladesh garment sector, the initiating force was private, with the government playing almost no role except for the fundamental one in India of providing good education.

The establishment of a science park in Taiwan, China, and legislation in China to allow special economic zones to attract FDI resulted from an initial government stimulus. A critical input for success was foreign participation that dealt with some of the roles cited above as components of industrial policy (source of new technology, facilitation of learning, source of new product ideas, centralized marketing allowing economies of scope, and coordination of entry of complementary firms). In China, the special economic zones mimicked the effects of a free trade policy, neutralizing adverse public policies. The zones did not discriminate among sectors. The decision by Taiwan, China, to foster a science park comes closer to a proactive industrial policy, but the experience at Hsinchu has not been systematically evaluated.

Many countries have attempted to use export-processing zones to attract FDI and perhaps generate agglomeration economies. Evaluation suggests that while potentially useful, they have had indifferent results (for references, see World Bank 2005). There have been a few success stories, such as the Republic of Korea and Taiwan, China, in the 1950s and early 1960s and the special economic zones of China. But there have been more than a thousand such efforts. There are few clues in the existing literature about why some export processing zones have been successful, while most have failed. (for a review and an evaluation of the Philippine experience, see Calanog 2006).

### *The Indian Software Industry*

In India a precondition for the development of the software industry was high-quality education in junior colleges and universities **financed** by the government. University graduates went abroad for further training, remained as expatriates in the **high-**technology sector, and later returned home or interacted intensively with newer **Indian firms**. The lamented brain drain became, with a lag, a source of strength and a critical catalytic input to development of the Indian software industry.

**Large numbers of English-trained programming graduates.** In the 1980s there were a growing number of programming graduates at levels ranging from postsecondary technical schools to those trained at the Indian Institutes of Technology, and many were underemployed. Almost all of them had been educated in English. The government's continuing investments in education had resulted in more than 1,800 educational institutions and polytechnics producing 70,000 to 85,000 computer science graduates every year (James 2000).<sup>12</sup> Many Indian graduates also had a second university degree or postgraduate degree from schools in the United States or the United Kingdom, often in computer technology (Deshmukh 1993). Other Indian software programmers received training in private software institutes to keep abreast of developments in the industry and acquired a breadth of software skills. Hence, many were familiar with main computer hardware systems (Lakha 1990), computer-aided software engineering tools, object-oriented programming, graphical user interface, and client networking (Lekshman and Lal 1998).

**Series of serendipitous events.** The main impetus to demand came from abroad from a series of serendipitous events. In the 1990s the ratio of world prices for programming services relative to those in India rose because of a global shortage of programmers and the demands for solutions to the anticipated Y2K problem. Enterprising businesses in India capitalized on this opportunity by setting up firms that were essentially employment agencies. Indian software programmers were hired on behalf of clients in the United States on short-term contracts to provide onsite services. "Bodyshopping," as this practice was called, became the predominant mode of Indian software exports. The development work was performed on the client's premises, saving software firms the high costs of acquiring computer hardware. The National Association of Software and Service Companies, the software trade association, reported that the software sector earned \$2.5 billion from Y2K billing from 1996 to 1999, a critical period in the growth of the industry (Software & Information Industry Association 2001). As late as 1988 software exports had been less than \$200 million. By 1998 they were \$3.6 billion, accounting for more than 10 percent of total Indian exports.

Indian software firms also benefited from another fortuitous event, the European Union's move to the euro. Many Indian software professionals were involved in adapting computer systems and databases to accommodate the euro. Between 2000 and 2002 India earned an estimated \$3 billion in revenues from euro-related information technology projects. A contributing factor was the level of programming costs in India, which conferred a Ricardian comparative advantage in some subsectors of software. As late as 1995, after substantial wage increases because of a rising demand for Indian software services, the annual wages of Indian software professionals were only 14–59 percent those of their counterparts in Canada, Switzerland, the United Kingdom, and the United States. This combination of skills and cost

savings led firms in some industrial countries to outsource their software development requirements to India.

Thus random events—the Y2K problem and the shift to the euro—exerted positive feedback and generated a succession of mutually reinforcing benefits. As for industrial policy, of whatever form, it seems unlikely that any government could have foreseen and acted on these serendipitous demands.

***The foreign role.*** A main contributor to development of the Indian software industry was the large number of expatriate Indian information technology professionals in Silicon Valley. In 1998, 774 (9 percent) of the high-technology firms were led by Indian chief executive officers (CEOs) (James 2000). Many Indian expatriates helped to convince large firms such as Oracle, Novell, and Bay Networks to establish operations in India (Saxenian 1999). Aware of the obstacles some Indians faced in raising capital for their software start-ups in India, they actively raised venture capital from U.S. investment firms and organized conferences in the United States to heighten awareness of the potential of India's software industry (Kripalani 2000). Finally, some of these expatriates lobbied the Indian government to revamp its telecommunication policies and other regulations that had impeded growth of the Indian software industry (Kripalani 2000).

FDI accounted for a large share of early investment in the sector—70 percent in Bangalore in 1996, for example (The Economist 1996). And this contribution understates the true impact. Texas Instruments, the first foreign firm to establish an offshore software facility in Bangalore in 1984, augmented Bangalore's inadequate land-based telecommunication infrastructure by investing in its own satellite communications network. Some of its lines were later leased to other software firms, enabling them to expand their India-based operations instead of relying solely on onsite services abroad. Until the government built software technology parks in the 1990s linked to earth stations and other telecommunications infrastructure, Texas Instrument's satellite network remained an important driving force behind the offshore development of software exports.

Once U.S.-based firms became interested in India, Bangalore's reputation for technical excellence and its abundant supply of information technology graduates made it a natural choice for foreign companies to locate their software business there (Stremlau 1996). With FDI came much of the infrastructure and international knowledge that allowed Indian firms to exploit international opening. Indian software firms also benefited from foreign joint ventures and partnerships, which created markets for Indian software exports. Partnerships with foreign firms added to the credibility of Indian firms, serving as an endorsement of its quality and reliability without government encouragement. Thus, other foreign firms looking to outsource their software development would invariably choose a software firm with a proven track record with another foreign company. And for small Indian firms attempting to move out of the low-end of the software business by venturing into software

packaging, having foreign partners gave them access to an established distribution network and knowledge of recent trends in the software market (because of proximity to demand in the United States). It also significantly lowered marketing costs. Because marketing costs account for as much as 70–80 percent of the final price of a software package (Lakha 1994), small Indian firms without a known brand, an extensive sales network, or sufficient revenue found it more profitable to sell their packages through a foreign collaborator.

***The Indian software industry and the new industrial policy.*** How does this experience of a successful sector square with the many strands of new industrial policy? All of it was privately initiated. Governments at various levels became involved only after the success of the sector was evident, ratifying the success rather than catalyzing it. The industry expanded on the basis of comparative advantage and never needed any protection. Indeed, one advantage of the software sector was that its inputs, largely downloads from satellites, and its output, uploaded to satellites, could not be easily taxed by the Indian authorities. A symbiosis of foreign and domestic firms was critical. Although there was clearly an agglomeration of firms in Bangalore, this was achieved spontaneously without government direction. Foreign contracts rather than government subsidies provided the basis for international exploration of markets. There is no evidence of government initiation or preference.

## Is Industrial Policy Still Relevant?

From Hamilton and List to contemporary discussions of industrial policy, the implicit framework has been that of a firm producing tradable goods at an initial cost disadvantage because of the limited industrial history of the country, learning to become more efficient, and then competing with imports in the local market or successfully exporting. Marketing of the efficiently manufactured product was implicitly assumed to be routine. Reduction of production costs, whether through internal learning by doing or through spillovers within industrial clusters, was viewed as paramount. In discussions of postwar Asian experience, some attention was given to the catalytic role of Japanese, Korean, and Taiwanese trading companies in assembling large quantities of goods and achieving scale economies in marketing, but this activity was not given center stage (Lall and Keesing 1992). Even if countries today could pursue the export-oriented policies of the Republic of Korea and Taiwan, China, of four decades ago, it is not clear that they would be efficacious, given the changed nature of both retailing and production networks.

In the last two decades there has been a shift in the institutional mechanism of international trade, as two types of organization have evolved. One is international production networks in which a producing firm organizes large numbers of suppliers

in several locations. The other is buyer-led networks, in which large retail chains provide specifications for the desired final product and encourage suppliers in developing countries to organize their own production system, which most often include large numbers of local subcontractors.<sup>13</sup> These networks have become increasingly important and are dominant in clothing and electronics and are growing in importance in products such as automotive components. In East Asia, in recent years, components "constitute at least a fifth of manufacturing exports and ... have typically grown 4–5 percent faster than overall trade in East Asia" (Yusuf and others 2003, p. 272).

One effect of the growing importance of international product networks is their efficiency at organizing production and continuously reducing costs so that the global price that nonmember firms must compete with shifts down rapidly. Infant firms undergoing learning face other hurdles: rapidly improving quality, changing characteristics of products, and an array of new goods that compete with existing ones (Ernst 2002). For firms attempting to enter export markets, it cannot be assumed that simply achieving low cost is sufficient to realize foreign sales. There is no guarantee that lead firms will be able to identify one or two firms in a small African country. The existence of supply networks imposes a significant challenge for developing-country firms that are not embedded in such a network, because the lead firms usually succeed in generating higher performance in design, engineering, effective use of information and communication technology, and ability to coordinate production in several locations (Yusuf and others 2003).

Further militating against the classical view of infant industries is the change in the nature of retailing. Consider a mundane product such as socks that can be produced efficiently with relatively labor-intensive technology. Huge retailers such as Wal-Mart and Target buy socks in quantities that exceed the production capacity of small (by international standards) companies. The special economic zones in China have become a series of clusters that produce enormous quantities of socks, ties, and other clothing. Retailers and wholesalers place large orders that are well beyond the production capacities of smaller firms, even if they have quickly learned to become cost-competitive in relatively small quantities. "These days buyers from New York to Tokyo want to be able to buy 500,000 pairs of socks all at once, or 300,000 neckties, 100,000 children's jackets" (Barboza 2004, section D, p. 1). European firms buy smaller, more varied products but expect local suppliers to provide "in-house design and sample making capabilities that would allow them to translate and adapt the design from Europe" (Sturgeon and Lester 2002, p. 49).

In textiles, clothing, electronics, automotive parts, and other sectors being part of an international product network is critical to exporting and quality upgrading. Firms that are not part of such networks may not succeed even if they are as efficient as members in production costs. Local participants in the network must "label, track, respond to product orders in real time on the basis of style, color, fabric, and

size; exchange information on an ... electronic basis, provide goods to a retailer's distribution center that can be efficiently moved to stores ... including containers with bar codes concerning contents" (Yusuf and others 2003, p. 283). These requirements, now fairly standard in many product areas, suggest that successful penetration of high-income markets will become increasingly difficult for countries that have not yet industrialized.

In electronics, an important labor-intensive growth sector in the past for many Asian countries, much of the production is now carried out by contract manufacturers that have grown enormously in the last decade. The activities of firms such as Solecron and Flextronics, formerly undertaken by main industrial country firms, are now outsourced. Sturgeon and Lester (2002) examined the location of several activities of Solecron—headquarters, manufacturing, materials purchasing and management, new production introduction centers, and after sales repair centers—and found that most of these activities take place in industrial countries or in the more advanced semi-industrial countries contiguous to them, such as Mexico, Puerto Rico, Romania, and Turkey. Ernst (2002, p. 24) confirms these results and points out that specialized clusters in countries such as the Nordic countries, France, Germany, and the United States are main sources as are Hungary, Israel, the Republic of Korea, Singapore, and Taiwan, China. Poorer countries even if they have a potential cost advantage after a long learning period will have trouble breaking into these existing networks.

Moreover, China and India present formidable competitors, as demonstrated by the concern over the termination of the Multifiber Arrangement and the Agreement on Textiles and Clothing and the potential losses incurred by countries that formerly had guaranteed access to Organisation for Economic Co-operation and Development markets. While it might be argued that the two giant countries will encounter rising wages and thus will enter more capital- and technology-intensive sectors, making room for new countries, both still have hundreds of millions of workers, largely in the rural sector, who remain poor and will keep a lid on real wages faced by industrialists over the next decades, implying a continuing supply of low-cost products in many sectors. While in principle poorer countries can find niches in which they have a comparative advantage, finding them is likely to require skills that are best nurtured by membership in a production network or direct interaction with large retailers.

What does the growing importance of production networks imply for potential government interventions? The Republic of Korea and Taiwan, China, had numerous trading companies that aggregated the orders of local manufacturers, following the Japanese model of the *shosha sogu*. Most of these arose spontaneously from private efforts. Governments could encourage the development of trading companies where there are market failures—setup costs may be high whereas the marginal costs of adding firms to the network may be small. Such trading firms would operate

across clusters of manufacturing firms. Again, this assumes that there are capital market failures that preclude a nascent trading firm from obtaining finance.

Other policy questions arise. Will government-sponsored clusters be as effective in generating continuing improvements in product development, quality upgrading, and efficiency to sustain competition on the world market or will firms within clusters improve faster by becoming part of networks? There is some anecdotal evidence that international networks attempt to limit the extent of upgrading, especially in higher value-added segments of design. If so, the question once again is whether to promote specific activities within the entire production nexus, but this is beyond the capacity of all but the most competent of governments (Humphrey and Schmitz 2002). Taiwan, China's experience in the Hsinchu Science Park may be an exception, but it is so far unconfirmed by systematic evidence.

## Concluding Remarks

Does the current policy landscape of the multilateral trading system even permit developing countries to pursue industrial policy? Should it? Developing countries have to contend with several multilateral agreements that were not in existence when the rich countries of today were developing. Have the constraints and disciplines imposed by World Trade Organization (WTO) agreements such as Trade-Related Aspects of Intellectual Property Rights, and Trade-Related Investment Measures become too restrictive to allow developing countries to chart their preferred course to economic development? This is a difficult question, but it cannot be dismissed out of hand. Certainly, the international policy environment today imposes constraints on the use of national policies that were absent even 15 years ago, and the constraints are backed by the potent dispute settlement procedure of the WTO (Noland and Pack 2003, chapter 5).

The experience in several countries in the last two decades suggests that private firms have often been successful in pursuing learning strategies that earlier analysts had advocated. The growth of the Indian software sector, Bangladesh's clothing industry, and China's special economic zones was driven primarily by private-sector agents (often from abroad). In Bangladesh and India the main role of the government was benign neglect, whereas China imitated the earlier success of Singapore by enabling the location of foreign investment in enclaves that were well provided with infrastructure. Much of the earlier investments came from overseas Chinese.

In none of these cases was there a government policy that singled out individual firms or industries with high learning potential and likely spillovers. In Bangladesh and China foreign firms brought standard technology but important extensive marketing networks. Standard comparative advantage can explain the pattern of sector choice. Compared with the exceptionally complex process of either picking sectors

(or firms) or allowing firms to identify their own competitive advantage, it seems much more efficient in the current state of intensifying world competition and the growing importance of extensive and complex supply networks to allow foreign firms to facilitate cost reduction in the host economy.

This suggests a change in focus from even the new industrial policy to one that focuses on negotiation with multinational firms on issues ranging from environmental regulation and taxes to efforts to ensure local learning. The difficulty with this approach is the limited amount of FDI going to developing countries—many countries in Africa, the Middle East, and Latin America continue to receive little. This may be because of their weak overall economic prospects given their poor policies. But in these economies hewing to some of the main tenets of the Washington Consensus (while recognizing some of its weaknesses) might prove a better investment of limited government competence and legitimacy than the extraordinarily complex strategies required by either the new or the old industrial policy.

## Notes

Howard Pack is a professor of business and public policy, economics, and management in the Wharton School at the University of Pennsylvania, Philadelphia; his e-mail address is [packh@wharton.upenn.edu](mailto:packh@wharton.upenn.edu). Kamal Saggi is a professor of economics in the Department of Economics at the Southern Methodist University; his e-mail address is [lsaggi@smu.edu](mailto:lsaggi@smu.edu). The authors acknowledge support from the U.K. Department for International Development project "Global Trade Architecture and Development." Wenying Choy provided superb research assistance.

1. Bardhan's model is in the spirit of the original learning by doing model of Arrow, which posited learning that occurred in the machine-producing sector. Some of the endogenous growth literature also posits such effects. However, the literature on technological innovation summarized in Evenson and Westphal (1995) and Ruttan (2001) shows that learning can occur in **all** sectors, a fact that would enormously complicate the results of much of the literature.

2. Paclr and Saggi (2001) explore the implications of the provision of free technology by the purchasers of a firm's exports, a further complication.

3. It is not likely that this criterion has been satisfied by the European Airbus effort, widely considered a main example of a successful industrial policy. Furthermore, one also needs to account for the cost of distortions that are generated by the taxes needed to finance the subsidies.

4. Succar's emphasis on the capital goods sector is similar in spirit to Arrow's learning by doing model and endogenous growth models such as Romer's (1986), which employed it as a building block.

5. Rodriguez-Clare (1996a) has shown that coordination failures can lead to "development traps."

6. As will be discussed below, good  $x$  could be produced by multinationals that establish local production, thus obviating the coordination problem.

7. Much of the effort of the Ministry of International Trade and Industry and the Ministry of Finance in Japan can be described as the interchange of information among firms and interaction with the government to reduce any obstacles to the realization of consistent plans. The same is true of French indicative planning of the 1950s and 1960s. As noted earlier, it is difficult to assess whether such sector-specific targeting was successful. For an extensive review of the empirical evidence on Japan, see Noland and Pack (2003).

8. The complexity of these issues is underlined by the fact that Bagwell and Staiger (1989) reach still other conclusions. They argue that if asymmetric information blocks the entry of high-quality firms,

export subsidies can improve welfare by breaking the entry barrier facing high-quality firms. Thus, whether an export subsidy is desirable hinges on the nature of the distortion that is caused by the presence of asymmetric information.

9. It is worth noting that if production is intended primarily for a protected domestic market, local suppliers, especially if there are local content requirements, may have costs above world prices, raising the possibility that greater linkages may lower the value of domestic output.

10. The following paragraphs are based on Noland and Pack (2003, chapter 2).

11. Japan's Ministry of Finance apparently agrees. In a June 2002 report issued by its Policy Research Institute, it maintains that "the Japanese model was not the source of Japanese competitiveness but the cause of our failure" and specifically argues that sectors sheltered by Ministry of International Trade and Industry became bloated and inefficient, whereas those exposed to international competition tended to be more market-aware, efficient, and profitable (Morita 2002).

12. Some observers feel this was an incorrect allocation of education funds and that the returns would have been greater to more extensive and higher quality primary and secondary education. The success of the software industry does not disprove this view. For example, the favorable effect of the adoption of the green revolution package on the income of Indian farmers of elementary school education is well established.

13. A good description of these alternatives and evidence on their quantitative importance is given in Gereffi (1999); see also Yusuf and others (2003, chapter 7). UNCTAD (2001) and Sturgeon and Lester (2002) provide evidence on the empirical importance of the international production networks.

## References

- Bagwell, Kyle, and Robert W. Staiger. 1989. "The Role of Export Subsidies When Product Quality is Unknown." *Journal of International Economics* 27(1-2):69-89.
- Baldwin, Robert E. 1969. "The Case against Infant Industry Protection." *Journal of Political Economy* 77(3):295-305.
- Barboza, David. 2004. "In Roaring China, Sweaters Are West of Soles City." *New York Times*. December 24.
- Bardhan, Pranab. 1971. "On Optimum Subsidy to a Learning Industry: An Aspect of the Theory of Infant-Industry Protection." *International Economic Review* 12(1):54-70.
- Batra, Geeta, and Hong W. Tan. 2002. *Inter-Firm Linkages and Productivity Growth in Malaysian Manufacturing*. International Finance Corporation, Washington, D.C.
- Beason, Richard, and David E. Weinstein. 1996. "Growth, Economies of Scale, and Targeting in Japan (1955-1990)." *Review of Economics and Statistics* 78(2):286-95.
- Besley, Timothy, and Anne Case. 1993. "Modeling Technology Adoption in Developing Countries." *American Economic Review* 83(2):396-402.
- Blalock, Garrick. 2001. "Technology from Foreign Direct Investment: Strategic Transfer Through Supply Chains." Cornell University, Department of Applied Economics and Management, Ithaca, N.Y.
- Calanog, Victor. 2006. "Essays on the Economics of Development Strategies." Ph.D. diss. The Wharton School, University of Pennsylvania, Philadelphia.
- Coase, Ronald. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3:1-44.
- Deshmukh, Vikas. 1993. "Bangalore: India's Hi-Tech Birthplace." *Economic Reform Today* 3.
- The Economist. 1996. "Software in India: Bangalore Bytes." *The Economist*. March 23.
- Ernst, Dieter. 2002. "Global Production Networks in East Asia's Electronics Industry." East-West Center, Honolulu, Hawaii.

- Ethier, Wilfred J. 1982. "National and International Returns to Scale in the Modern Theory of International Trade." *American Economic Review* 72(3):389–405.
- Evenett, Simon. 2003. "Study on Issues Related to a Possible Multilateral Framework on Competition Policy." WT/WGTCP/W228. World Trade Organization, Geneva.
- 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 3a*. Amsterdam: North Holland.
- Gereffi, Gary. 1999. "International Trade and Industrial Upgrading in the Apparel Commodity Chain." *Journal of International Economics* 48(1):37–70.
- Gereffi, Gary, and Olga Memedovic. 2003. "The Global Apparel Value Chain: What Prospects for Upgrading by Developing Countries?" Sectoral Studies Series 12770. United Nations Industrial Development Organization. Vienna.
- Grossman, Gene, and Henrik Horn. 1988. "Infant Industry Protection Reconsidered: The Case of Informational Barriers to Entry." *Quarterly Journal of Economics* 103(4):767–87.
- Hausmann, Ricardo, and Dani Rodrik. 2003. "Economic Development as Self-Discovery." *Journal of Development Economics* 72(2):603–33.
- Hobday, Mike. 1995. *Innovation in East Asia: The Challenge to Japan*. London: Edward Elgar.
- Hoff, Karla. 1997. "Bayesian Learning in an Infant Industry Model." *Journal of International Economics* 43(3–4):409–36.
- Huang, Yasheng. 2003. *Selling China*. Cambridge: Cambridge University Press.
- Humphrey, John, and Hubert Schmitz, 2002. "Governance and Upgrading: Linlring Industrial Cluster and Global Value Change Research." Working Paper 120. University of Sussex, Institute for Development Studies, Brighton. United Kingdom.
- James, David. 2000. "India Starts Up." *Upside Magazine* 12(4):265–69.
- Javorcik, Beata S. 2004. "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages." *American Economic Review* 94(3):605–27.
- Kim, Linsu. 1997. *Imitation to Innovation: The Dynamics of Korea's Technological Learning*. Boston: Harvard Business School Press.
- Klimenko, Mikhail. 2004. "Industrial Targeting, Experimentation and Long-run Specialization." *Journal of Development Economics* 73(1):75–105.
- Kripalani, Manjeet. 2000. "A Typhoon of Venture Capital?" *Business Week*. January 31.
- Laderman, Daniel, William F. Maloney, and Luis Servén. 2003. *Lessons from NAFTA for Latin American and Caribbean Countries*. Washington, D.C.: World Bank.
- Lakha, Sali. 1990. "Growth of the Computer Software Industry." *Economic and Political Weekly*, January 6:49–56.
- . 1994. "The New International Division of Labour and the Indian Computer Software Industry." *Modern Asian Studies* 28(2):381–408.
- Lall, Sanjaya, and Donald Keesing. 1992. "Marketing Manufactured Exports from Developing Countries: Learning Sequences and Public Support." In G. Helleiner, ed., *Trade Policy, Industrialization, and Development: New Perspectives*. Oxford: Clarendon Press.
- Larraz B., Felipe, Luis F. Lopez-Calva, and Andres Rodriguez-Clare. 2000. "Intel: A Case Study of FDI in Central America." CID Working Paper 58. Harvard University, Center for International Development, Cambridge, Mass.
- Lawrence, Robert Z., and David E. Weinstein. 2001. "Trade and Growth: Import Led or Export Led? Evidence from Japan and Korea." In Joseph E. Stiglitz and Shahid Yusuf, eds., *Rethinking the East Asian Miracle*. Oxford: Oxford University Press.

- Lekshman, Pai, and Uday Lal. 1998. "A Major Exporter Turns Inward." *Computerworld* 32(14), April 6.
- Lee, Jong-wha. 1996. "Government Interventions and Productivity Growth in Korean Manufacturing Industries." *Journal of Economic Growth* 1(3):391–414.
- Markusen, James R., and Anthony Venables. 1999. "Foreign Direct Investment as a Catalyst for Industrial Development." *European Economic Review* 43(2):335–56.
- Mayer, Wolfgang. 1984. "The Infant-Export Industry Argument." *Canadian Journal of Economics* 17(2):249–69.
- McKinsey Global Institute. 1998a. *Productivity — The Key to an Accelerated Development Path for Brazil*. Washington, D.C., and Sao Paulo: McKinsey & Co.
- . 1998b. *Productivity Led Growth in Korea*. Washington, D.C. and Seoul: McKinsey & Co.
- Moran, Theodore. 1998. *Foreign Direct Investment and Development*. Washington, D.C.: Institute for International Economics.
- . 2001. *Parental Supervision: The New Paradigm for Foreign Direct Investment and Development*. Washington, D.C.: Institute for International Economics.
- Morita, Issei. 2002. *Financial Times*. June 27.
- Noland, Marcus, and Howard Pack. 2003. *Industrial Policy in an Era of Globalization: Lessons from Asia*. Washington, D.C.: Institute for International Economics.
- Okuno-Fujiwara, Masahiko. 1988. "Interdependence of Industries, Coordination Failure, and Strategic Promotion of an Industry." *Journal of International Economics* 25(1/2):25–43.
- Pack, Howard. 2000. "Industrial Policy: Growth Elixir or Poison?" *World Bank Research Observer* 15(1):47–68.
- Pack, Howard, and Larry Westphal. 1986. "Industrial Strategy and Technological Change: Theory versus Reality." *Journal of Development Economics* 22(1):87–128.
- Pack, Howard, and Kamal Saggi. 2001. "Vertical Technology Transfer via International Outsourcing." *Journal of Development Economics* 65(2):389–415.
- Rhee, Yung W. 1990. "The Catalyst Model of Development: Lessons from Bangladesh's Success with Garment Exports." *World Development* 18(2):333–46.
- Rodriguez-Clare, Andres. 1996a. "The Division of Labor and Economic Development." *Journal of Development Economics* 49(1):3–32.
- . 1996b. "Multinationals, Linkages, and Economic Development." *American Economic Review* 86(4):852–73.
- . 2004a. "Clusters and Comparative Advantage: Implications for Industrial Policy." Research Department Working Paper 523. Inter-American Development Bank, Washington, D.C.
- . 2004b. "Microeconomic Interventions After the Washington Consensus." Research Department Working Paper 524. Inter-American Development Bank, Washington, D.C.
- Rodrik, Dani. 1996. "Coordination Failures and Government Policy: A Model with Applications to East Asia and Eastern Europe." *Journal of International Economics* 40(1–2):1–22.
- . 2004. "Industrial Policy for the Twenty-First Century." CEPR discussion Paper 4767. Centre for Economic Policy Research, London.
- Romer, Paul M. 1986. "Increasing Returns and Long Run Growth." *Journal of Political Economy* 94(5):1002–37.
- Rosen, Daniel H. 1999. *Behind the Open Door: Foreign Enterprises in the Chinese Marketplace*. Washington, D.C.: Institute for International Economics.
- Ruttan, Vernon. 2001. *Technology, Growth, and Development: An Induced Innovation Perspective*. New York: Oxford University Press.

- Saggi, Kamal. 2002. "Trade, Foreign Direct Investment, and International Technology Transfer: A Survey." *World Bank Research Observer* 17(2):191–235.
- Saxenian, AnnaLee. 1999. *Silicon Valley's New Immigrant Entrepreneurs*. San Francisco: Public Policy Institute of California.
- . 2001. "Taiwan's Hsinchu Region: Imitator and Partner for Silicon Valley." Policy Paper 00–44. Stanford Institute for Economic Policy Research, Stanford, Calif.
- Scitovsky, Tibor. 1954. "Two Concepts of External Economies." *Journal of Political Economy* 62(2):143–51.
- Software & Information Industry Association. 2001. "Submission of the Software & Information Industry Association to the U.S. Trade Representatives: 2000 Special 301 Recommendations, February 18, 2000." (March 18, 2001). Washington, D.C.
- Stremlau, John. 1996. "Dateline Bangalore: Third World Technopolis." *Foreign Policy* 102:152–59.
- Sturgeon, Timothy, and Richard Lester. 2002. "Upgrading East Asian Industries: New Challenges for Local Suppliers." Massachusetts Institute of Technology, Industrial Performance Center, Cambridge, Mass.
- Sturgeon, Timothy, and Richard Lester. 2003. "The New Global Supply-Base: New Challenges for Local Suppliers in East Asia." Working Paper 03-001. Massachusetts Institute of Technology, Industrial Performance Center, Cambridge, Mass.
- Succar, Patricia. 1987. "The Need for Industrial Policy in LDCs—A Restatement of the Infant Industry Argument." *International Economic Review* 28(2):521–34.
- UNCTAD (United Nations Conference on Trade and Development). 2001. *World Investment Report: Promoting Linkages*. New York.
- Williamson, John. 1990. "What Washington Means by Policy Reform." In John Williamson, ed., *Latin American Adjustment: How Much Has Happened?* Washington, D.C.: Institute for International Economics.
- World Bank. 1993. *The East Asian Miracle*. Washington, D.C.: The World Bank.
- . 2005. *World Development Report 2005: A Better Investment Climate for Everyone*. Washington, D.C.: World Bank and Oxford University Press.
- . 2006. *Doing Business in 2006*. Washington, D.C.: World Bank.
- Yusuf, Shahid, M. Anjum Altaf, Barry Eichengreen, Sudarshan Gooptu, Kaoru Nabeshiia, Charles Kenny, Dwight H. Perkins, and Marc Shotten. 2003. *Innovative East Asia: The Future of Growth*. Washington, D.C.: World Bank and Oxford University Press.

Economics from Oxford

OXFORD  
UNIVERSITY PRESS



# World Development Report 2006 *Equity and Development* World Bank

*The World Development Report 2006: Equity and Opportunity* presents a social development strategy organized around the themes of social inclusion, cohesion, and accountability. It examines equality of opportunities—a potentially important factor affecting both the workings of the investment environment and the empowerment of the poor—by building on and extending existing accountability frameworks presented in the 2005 Report.

Now in its twenty-eighth edition, the World Development Report offers practical insights for policymakers, business developers, economic advisers, researchers, and professionals in the media and in non-governmental organizations. It is also an essential supplement to economic and development courses in both academic and professional settings.

January 2006

Paperback | 0-8213-6249-6 | £15.99 | \$26.00

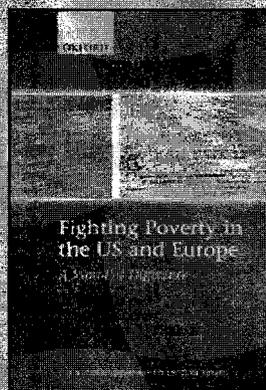
Hardback | 0-8213-6251-8 | £29.99 | \$50.00

For more information and ordering details,  
please visit [www.oup.co.uk/socsci](http://www.oup.co.uk/socsci)

Economics from Oxford

OXFORD  
UNIVERSITY PRESS

NEW IN PAPERBACK



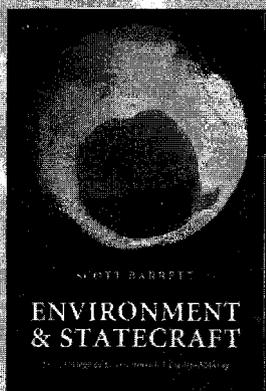
## Fighting Poverty in the US and Europe

### *A World of Difference*

**Alberto Alesina & Edward Glaeser, both at Harvard University.**

- Concise, multidisciplinary, and timely study of the differences between the US and Europe's welfare systems
- Written by top-notch economists drawing on in-depth original research

November 2005 | Paperback | 0-19-928610-8 | £12.99 | \$24.95



## Environment and Statecraft

### *The Strategy of Environmental Treaty-Making*

**Scott Barrett, Johns Hopkins University.**

- Develops a coherent and consistent theory that integrates different approaches to the study of international cooperation and the environment
- Provides information on over 300 treaties

November 2005 | Paperback | 0-19-928609-4 | £12.99 | \$24.95

For more information and ordering details,  
please visit [www.oup.co.uk/socsci](http://www.oup.co.uk/socsci)

# Stay alert to the very latest research

Design your own email alerts to track journal content that is important to you.

CiteTrack will alert you by email whenever new content is published that matches criteria based on the topics, authors and articles you want to track.

#### Topics

Tell CiteTrack which words or subjects to watch for in new content.

#### Authors

Be alerted whenever key authors you are following publish a new paper.

#### Articles

Know whenever a paper of interest to you is referenced by another paper or letter.

CiteTrack checks for content matching your criteria in new issues of the journal you've asked it to monitor, and emails you each time new content is found.

You will receive an email every time matching content is found in new issues, one email for each new issue of each journal you are tracking.

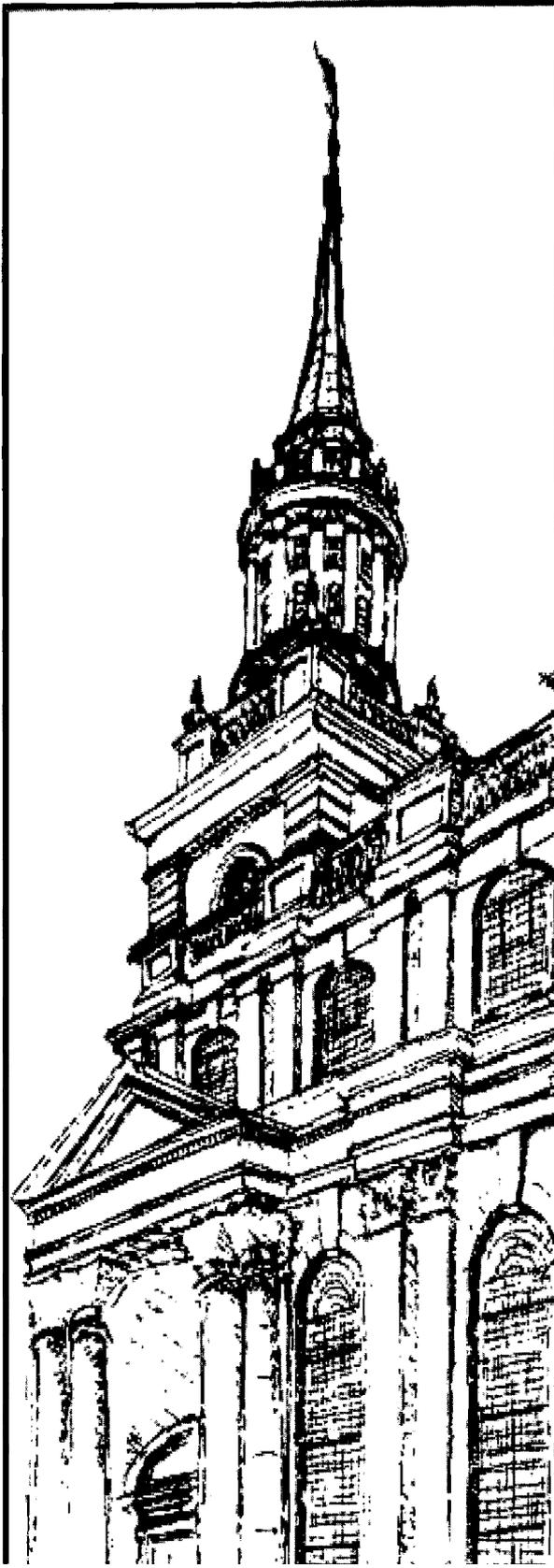
With over 340 journals currently available through this service, you can easily monitor a wealth of quality research with one simple service.

Sign up for CiteTrack today – go to [www.highwire.org](http://www.highwire.org), and click on 'My Email Alerts' for more information.



[www.oupjournals.org](http://www.oupjournals.org)

**OXFORD**  
UNIVERSITY PRESS



# Oxford

## Journals Advertising

For information on  
advertising in this journal  
please contact:

Helen Pearson  
Oxford Journals Advertising  
PO Box 347  
Abingdon, OX14 1GJ, UK

Tel: +44 (0) 1235 201904

Fax: +44 (0) 8704 296864

e-mail: [helen@oxfordads.com](mailto:helen@oxfordads.com)

**THE WORLD BANK**

1818 H Street NW

Washington, DC 20433, USA

World Wide Web: <http://www.worldbank.org/>

E-mail: [researchobserver@worldbank.org](mailto:researchobserver@worldbank.org)

