

Introduction Boris Pleskovic and Nicholas Stern

23244
July 2001

Annual

Keynote Addresses

Mats Karlsson
Joseph E. Stiglitz
Jeffrey Sachs
János Kornai

World Bank

New Development Thinking

Paul Collier
Dani Rodrik
Jan Willem Gunning
Karla Hoff

Conference

Crises and Recovery

William Easterly, Roumeen Islam, and Joseph E. Stiglitz
Ricardo J. Caballero and Mohamad L. Hammour
Eisuke Sakakibara
Daniel Lederman, Ana Maria Menéndez,
Guillermo Perry, and Joseph E. Stiglitz

on Development

Corporate Governance and Restructuring

I. J. Alexander Dyck
Gérard Roland

Economics

Social Security, Public and Private Savings

J. Michael Orszag and Peter R. Orszag
Orazio P. Attanasio and Miguel Székely



2000

Edited by **Boris Pleskovic and Nicholas Stern**

Annual
World Bank
Conference
on Development
Economics
2000

Edited by **Boris Pleskovic and Nicholas Stern**

The World Bank Washington, D.C.

© 2001 The International Bank for Reconstruction and Development / THE WORLD BANK
1818 H Street NW, Washington, DC 20433, U.S.A.
All rights reserved. Manufactured in the United States of America.

First printing July 2001
1 2 3 4 04 03 02 01

The Annual World Bank Conference on Development Economics is a forum for discussion and debate of important policy issues facing developing countries. The conferences emphasize the contribution that empirical and basic economic research can make to understanding development processes and to formulating sound development policies. Conference papers are written by researchers in and outside the World Bank. The conference series was started in 1989.

Conference papers are reviewed by the editors and are also subject to internal and external peer review. Some papers were revised after the conference, sometimes to reflect the comments by discussants or from the floor. Discussants' comments were not revised. As a result, discussants' comments may refer to elements of the paper that no longer exist in their original form. Participants' affiliations identified in this volume are as of the time of the conference, April 28–30, 1999.

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

The material in this publication is copyrighted. The World Bank encourages dissemination of its work and will normally grant permission promptly. Permission to photocopy items for internal or personal use, for the internal or personal use of specific clients, or for educational classroom use is granted by the World Bank, provided that the appropriate fee is paid directly to Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, U.S.A., telephone 978-750-8400, fax 978-750-4470. Please contact the Copyright Clearance Center before photocopying items.

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

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

Meta de Coquereaumont, Paul Holtz, Barbara Karni, and Alison Strong edited this volume. Jessica Saval prepared the cover and layout, and Molly Lohman managed editing and production. All are with Communications Development.

Annual World Bank Conference on Development Economics is indexed in *Human Resources Abstracts*, the *Index of Economic Articles*, the *Index of Social Science and Humanities Proceedings*, the *Index to International Statistics*, the *Public Affairs Information Service*, *Sage Public Administration Abstracts*, the *Standard Periodical Directory*, and *Ulrich's International Periodicals Directory* and online by ABI/INFORM and DIALOG.

ISSN 1020-4407
ISBN 0-8213-4981-3

Contents

Introduction	1
<i>Boris Pleskovic and Nicholas Stern</i>	
OPENING ADDRESS	7
<i>Mats Karlsson</i>	
KEYNOTE ADDRESSES	13
Development Thinking at the Millenium	
<i>Joseph E. Stiglitz</i>	
A New Global Consensus on Helping the Poorest of the Poor	39
<i>Jeffrey Sachs</i>	
Ten Years After <i>The Road to a Free Economy</i> : The Author's Self-Evaluation	49
<i>János Kornai</i>	
New Development Thinking	
Consensus Building, Knowledge, and Conditionality	67
<i>Paul Collier</i>	
Development Strategies for the 21st Century	85
<i>Dani Rodrik</i>	
Comments	
<i>Nancy Birdsall</i>	109
<i>François Bourguignon</i>	117
<i>Fukunari Kimura</i>	123
Rethinking Aid	125
<i>Jan Willem Gunning</i>	
Beyond Rosenstein-Rodan: The Modern Theory of Coordination Problems in Development	145
<i>Karla Hoff</i>	

Comments	
<i>Gustav Ranis</i>	177
<i>Charles C. Soludo</i>	183
Crises and Recovery	
Shaken and Stirred: Explaining Growth Volatility	191
<i>William Easterly, Roumeen Islam, and Joseph E. Stiglitz</i>	
Creative Destruction and Development: Institutions, Crises, and Restructuring	213
<i>Ricardo J. Caballero and Mohamad L. Hammour</i>	
Comment	
<i>Abhijit V. Banerjee</i>	237
The East Asian Crisis—Two Years Later	243
<i>Eisuke Sakakibara</i>	
Comment	
<i>Andrew Berg</i>	257
Mexico—Five Years After the Crisis	263
<i>Daniel Lederman, Ana María Menéndez, Guillermo Perry, and Joseph E. Stiglitz</i>	
Comment	
<i>Alejandro Werner</i>	283
Corporate Governance and Restructuring—Lessons from Transition and Crises	
Ownership Structure, Legal Protections, and Corporate Governance	291
<i>I. J. Alexander Dyck</i>	
Corporate Governance and Restructuring: Lessons from Transition Economies	331
<i>Gérard Roland</i>	
Comments	
<i>John S. Earle</i>	353
<i>Rafael La Porta</i>	361
Social Security, Public and Private Savings	
The Benefits of Flexible Funding: Implications for Pension Reform in an Uncertain World	373
<i>J. Michael Orszag and Peter R. Orszag</i>	
Household Saving in East Asia and Latin America: Inequality, Demographics, and All That	393
<i>Orazio P. Attanasio and Miguel Székely</i>	
Comment	
<i>Robert Holzmann</i>	439

Introduction

Boris Pleskovic and Nicholas Stern

The Annual World Bank Conference on Development Economics is a global gathering of scholars and practitioners of development policy from academic life, government, and the private sector. The topics selected for the conference vary from year to year, but we seek to include new areas of concern and cutting edge research, together with areas we believe can benefit from exposure to recent knowledge and experience.

The 12th annual conference, held at the World Bank on April 18–20, 2000, focused primarily on four areas: new development thinking, crises and recovery, corporate governance and restructuring, and social security, public and private savings. In addition, there were three keynote addresses: “Development Thinking at the Millennium” by Joseph E. Stiglitz, “A New Global Consensus on Helping the Poorest of the Poor” by Jeffrey Sachs, and “Ten Years After *The Road to a Free Economy*: The Author’s Self-Evaluation” by János Kornai.

Mats Karlsson in his opening address outlines a number of challenges for development. These include the intransigence of poverty in Africa, exacerbated by the ineffectiveness of the state and the widening AIDS crisis, and finding new ways of working with the private sector by establishing partnerships at the country and global levels.

Joseph E. Stiglitz, in his keynote address, identifies equilibrium and change as the twin foci of development economics. In describing and analyzing these two forces, the central question is, how do we know what we know or come to believe what we come to believe? Stiglitz’s response? A blend of nature and nurture. As economists we believe that incentives play a central role. Thus as countries make decisions on the best available evidence, he encourages a healthy skepticism when weighing the merits of various sources of policy advice. Inevitably, both internal and external policy advice is driven by assorted incentive structures. Over time there has been a noticeable change in thinking about development economics. Experience has shown that there is more to development than privatization, liberalization, and stabilization. Long-term sustainable growth requires development of a consensus behind the reform policies; they cannot be

Boris Pleskovic is research administrator at the World Bank. Nicholas Stern is senior vice president, Development Economics, and chief economist at the World Bank.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

imposed from the outside, which is part of the reason for the widespread failure of aid conditionality. He suggests that, contrary to the neoclassical model, institutions do matter, but so does the distribution of income. An alternative paradigm would include the role of imperfect and incomplete information. He favors a more comprehensive framework, including a greater sensitivity to the complexity of the development process. We should also show a greater sense of optimism about what the future might bring.

Jeffrey Sachs sets the tone for the discussion of crises and recovery in his keynote address on helping the poor, calling for a new consensus on economic development, to replace inadequate approaches to global poverty. He refers to the general failure to address the AIDS epidemic, malaria, and poverty, especially in the poorest parts of the world—Sub-Saharan Africa, parts of the Andean region, the Gangetic valley of India, and central Asia. He suggests that this failure is due primarily to inadequate funding and inappropriate structures to address the problem. The inadequate funding is due to lack of contributions from rich countries, especially the United States, while the International Monetary Fund (IMF) and, to a lesser extent, the World Bank are placing too much emphasis on economic reform and good governance, to the neglect of other key aspects. Sachs suggests that escape from poverty be based on economic reform, funding to address health and education needs, technology development, and structural adjustment, especially export diversification—which would require an industrial strategy and access to U.S. and European markets. To achieve these ends, he suggests a bipartisan approach. This would include many of the elements proposed by the Meltzer Commission. The IMF would get out of the poverty business and the World Bank would focus on the poorest countries. The World Bank should retreat from banking and refocus on knowledge creation.

The third keynote address, by **János Kornai**, provides a perspective on the recent history of transition economies. Drawing on experience in the Czech Republic, Hungary, Poland, and Russia, Kornai focuses on two issues: ownership reform with private sector development and macroeconomic stabilization. He argues that the choice between gradualism and shock therapy is a false dichotomy. While speed is important, it is not the primary indicator of success. Transformation of the economy and the society often proceeds in parallel, with many interactions. Thus it is essential to create favorable conditions for bottom-up development of the private sector. The shock tactics employed in the Czech Republic initially yielded some favorable political outcomes. They were less successful, however, looked at over a 10-year period than the more gradual reforms in Hungary and Poland. The shock approach in Russia was also a failure. One measure of success is change in labor productivity: over 10 years Hungary had gains of 36 percent and Poland of 29 percent, while the Czech Republic had a gain of 6 percent and Russia had a decline of 33 percent. He suggests emphasizing consolidation and stability, making growth sustainable, and not being overly concerned with speed.

New Development Thinking

Paul Collier argues that poverty reduction on a grand scale is now possible. He suggests this goal can be achieved if governments and donors can establish the policy

and institutional changes needed for broad-based growth. He attributes the older, pessimistic analysis of mass poverty to an undeservedly hostile response to progrowth policies and the failure to make such policies sufficiently broad based. Providing information and building consensus need to be treated as functional parts of the business of development. Attempts by donors to reduce poverty have met with only limited success, for a number of reasons. Like Stiglitz, Collier is skeptical that conditionality imposed from outside can succeed. Flexibility is needed to determine the most appropriate institutional design for each country. Governments should initiate institutional experiments and competitions to discover which institutions work best. A social consensus behind these experiments would speed these processes and enhance their credibility.

Dani Rodrik, in exploring development strategies for the 21st century, emphasizes the importance of markets underpinned by solid public institutions. While the idea of a mixed economy is a valuable heritage of the 20th century, the key lesson is that today's advanced industrial countries owe much of their success to having evolved their own workable models of a mixed economy. He places considerable importance on ownership of reform. Institutions are essential to protect property rights, regulate market participants, maintain macroeconomic stability, provide social stability, and manage conflict. These institutions can take many forms. They can also continue to evolve and to benefit from useful precedents in other societies. Developing countries should use available information to adapt institutional innovation to suit their own local practices and needs and not place undue emphasis on the global economy and the blueprints that emanate from it.

Jan Willem Gunning reviews the recent debate on aid's role in three areas: providing finance, changing policies in recipient countries, and transmitting knowledge. Providing finance, he argues, is a questionable way of accelerating growth rates except in a good policy environment. However, aid could play an insurance role in moderating the impact of negative external shocks on growth. As do other participants, Gunning argues that evidence confirms that aid cannot achieve sustainable policy reform. In particular, he feels that donor promotion of tax efforts in poor countries may be misguided. Rather, aid might be used for tax relief to avoid the high cost of taxation in the early growth stages. He proposes that donors consider *ex post* conditionality. He also suggests that transfers of knowledge need not be bundled with the provision of aid.

In the final article on new development thinking, **Karla Hoff** examines coordination problems in development. She offers a historical overview of how development economics has dealt with spillovers (externalities) over the past 30 years. Modern economic theory has broadened our view of the sources of spillovers that could lead to underdevelopment as an equilibrium. A critical determinant of action is the environment, including the behavior of other agents in that environment. Externalities that matter for welfare are not just direct interdependencies but rather many classes of systemic externalities. Initial differences in circumstances or beliefs may not only persist, but may be magnified over time. The basic lesson is that coordination failures abound, and neither the market alone nor government alone can solve them.

Crises and Recovery

William Easterly, Roumeen Islam, and Joseph E. Stiglitz look at volatility and growth, providing fresh insights into their relationship. They suggest that in explaining this relation traditional factors such as wage and price rigidities may have been overemphasized and factors relating to the financial system have been underemphasized. While recognizing the potential impact of a variety of microeconomic variables that could not be included in their analysis, the authors suggest that their findings, if correct, would have strong policy implications. Real wage flexibility may have little effect on volatility since the lowering of real wages may be offset by adverse demand effects. It is not evident that opening the capital account will help stability. They suggest the need to devise strategies to hedge against the risks of sudden outflows. Their analysis confirms that financial institutions play a central role in volatility—financial depth reduces volatility up to a point. They also note that too much private credit can increase volatility.

Ricardo J. Caballero and Mohamad L. Hammour examine a different aspect of the effect of crises on development. They focus on what Schumpeter termed “creative destruction.” This was considered to be at the core of economic growth, as it provided the means for technology to combine with opportunities to upgrade the production structure. An index of restructuring can be developed based on a variety of reallocation measures. One of the more successful reallocation measures is labor reallocation. The authors provide a theoretical framework, backed by some limited empirical work, to focus on potential obstacles to creative destruction. They find that the adverse effect of crises on restructuring is even costlier than the immediate impact on unemployment and other aggregate indicators might suggest. This freezing of the restructuring process and resulting stagnation in productivity comes primarily from financial constraints. Crises are also likely to result in a significant amount of liquidations that may be privately inefficient, leading to large job losses and liquidations of organizational capital.

Eisuke Sakakibara outlines the circumstances surrounding the East Asian crisis and then presents some of the choices warranting consideration by policymakers, if a repeat crisis is to be avoided. Perplexingly, macroeconomic indicators were strong among East Asian countries hit by the crisis. Market rate spreads and country risk ratings did not suggest macroeconomic weakness. At the same time many of the structural weaknesses in corporate and national governance had existed for decades. The immediate cause of the crisis can be attributed to a number of factors: liberalizing a weak banking sector and at the same time opening capital accounts, amassing aggregate short-term debt in excess of foreign reserves, and defending unsustainable pegs. While it is important for countries to strengthen their financial systems, the author argues that there is also a need for a lender of last resort to help prevent future currency crises. Since the current political situation precludes having a single institutional lender with free capital movement, he suggests an Asian Monetary Fund as a second-best solution. This would be preceded by cooperation in the real sector, including on free trade and direct investment. In the meantime policymakers could focus on developing well functioning capital markets. The

Japanese government could engage in regionwide efforts to stimulate primary markets for international bonds and upgrade the regional secondary markets.

Daniel Lederman, Ana María Menéndez, Guillermo Perry, and Joseph E. Stiglitz, in exploring the Mexico crisis and its aftermath, find a number of salient features similar to those in the East Asia crisis. Their focus is on the real economy. They conclude that a sharp depreciation of the currency in 1995 resulted in a precipitous fall in fixed investment as a result of a negative income effect and higher cost of capital. The authors also note the important role of credit availability, both in quantity and cost. A third of loans to Mexican firms were denominated in foreign currency, and most had been extended to firms without sources of foreign income. This availability of dollar loans together with the change in relative prices resulted in a shift toward tradable goods and a sharp rise in exports. In Mexico, as in Indonesia, the Republic of Korea, and Thailand, fixed investment growth fell during the crisis. The authors attribute the decline and subsequent rise in GDP growth primarily to the behavior of fixed investment. Depreciation of the currency can have adverse consequences for investment in the short run as a result of the effect on the relative price of capital. However, they argue that in the medium term a real depreciation is healthy because it stimulates growth in exports, increases the share of the tradable sector, and has a large multiplier effect on investment.

Corporate Governance and Restructuring: Lessons from Transition and Crises

I. J. Alexander Dyck provides a functional framework rather than an institutional or neoclassical economic perspective for examining governance. He draws on a wide variety of theoretical and empirical evidence to study the effect on governance of legal protections and the ownership structure of firms. The primary function of corporate governance, he argues, is to facilitate investment. Among the many findings the evidence suggests that the extent of legal protections for financiers is correlated with the depth of equity markets and the rate of initial public offering activity. Corporate governance is a major concern for policymakers. When it is not functioning effectively, resources for new projects dry up, slowing national growth and development. However, reforms that focus only on legal rules have limits. The effectiveness of legal protections depends on complementary institutions that provide enforcement and information for investors.

Gérard Roland also looks at the role of corporate governance, but he focuses on transition economies. He emphasizes the implications for efficiency, the heterogeneity of managers' skills, the diversity of firms' restructuring tasks and financial situations, and political constraints. He argues that corporate governance is even more important in transition economies, where the initial situation was state governance and the approach chosen will affect overall economic performance. He generally favors more gradual privatization. Under mass privatization incumbent managers get control, and this can lead to asset stripping. It may also lead to a sudden and strong concentration of economic power in the hands of insider managers.

Gradual sales, by contrast, yield a lower initial concentration of economic power, produce new entrepreneurs, and promote the emergence of a strong middle class. This in turn helps foster further reform and stable democracy.

Social Security, Public and Private Savings

J. Michael Orszag and Peter R. Orszag examine the implications of flexible funding for pension reform in an uncertain world. How much pressure pensions will exert on government budgets in the future is unclear because of substantial uncertainty surrounding estimates of pension costs over coming decades. The authors use a simplified model to analyze several questions. What are the likely implications of uncertainty? How should reform be undertaken, especially decisions on prefunding? They argue for the importance of taking into account not only best guesses about the future but also best assessments of the uncertainty surrounding these guesses. Changing circumstances may warrant a change in the degree of funding. They note that achieving some degree of reversibility may be easier in a public approach to prefunding. It may therefore be unwise to commit to the up-front costs associated with moving from a pay as you go to a prefunded system without any assurance that the costs can be recaptured if the reform turns out to be less beneficial than expected.

Orazio P. Attanasio and Miguel Székely explore household savings behavior at the microeconomic level in Mexico, Peru, Taiwan (China), and Thailand. They analyzed disaggregated groups by education level to see the relation of savings behavior to income distribution. They also test the relevance of the life-cycle model for explaining differences in savings behavior. They find large differences across economies and across education groups within each economy. East Asian households, especially younger households, have had greater capacity to save because of higher income growth, lower fertility rates, and a more advanced demographic transition. Household savings are generated by a wide range of households in East Asia, whereas in Latin America almost all household savings are generated by the richest 20 percent of the population. The authors find no strong evidence of negative saving in the last part of the life cycle in any of the four economies. Their results suggest that the demographic trends now being projected are unlikely by themselves to generate large increases in savings rates under current circumstances.

* * *

As in previous years, the planning and organization of the 2000 conference was a joint effort. Special thanks are due to Joseph E. Stiglitz for overall guidance. We also wish to thank staff members, especially F. Desmond McCarthy, and several anonymous reviewers for their assistance. And conference coordinators Julee Allen, Mantejwinder Jandu, and Jean Gray Ponchamni, whose outstanding organizational skills helped ensure a successful conference. Finally, we thank the editorial staff for pulling this volume together, especially Meta de Coquereumont, Paul Holtz, Barbara Karni, Molly Lohman, and Alison Strong, all of Communications Development Incorporated.

Opening Address

Mats Karlsson

Thank you, colleagues. I am pleased to be attending this conference and to have been asked to make the opening address. I recently joined the World Bank, but my real home—my community—is the same as yours: the community of development economists.

In the late 1970s, inspired by Amartya Sen, I studied philosophy and economics. But I was soon drafted into the aid business and moved to Africa, where I spent years struggling to make sense of structural adjustment. I argued for structural adjustment in terms of its political economy, as a means of fighting corruption and fighting for social justice. And I guess that is what I am still doing.

I would not be doing this job if I did not believe that development economics addresses humankind's most important challenges. Through it we face key questions on poverty and prices—questions that often have excruciatingly difficult answers amid a frustrating political reality. To succeed, we must be clear about the values that drive us. Economics is a powerful science, and I know that many of you have been prime movers inside and outside the Bank. But with that power comes responsibility.

Over the past few days the Bank was besieged by protestors. It was intellectually challenging to hear, try to make sense of, and debate some of what the protestors said. Not that there was much intellectual rigor in what was being said on the streets, but on television or in an auditorium of restless students, important questions were posed. Why, after so many years and so much money, have we failed to reduce poverty in Africa? If economic reforms often fail to deliver or even to be sustained, why can't we think of better approaches? How important to economic performance are democracy and participation? We have all heard these types of questions, and we need to come up with better answers.

Interestingly, the demonstrators might have done us a favor. With so many attacks on the Bank the demonstrators have focused many commentators on development

Mats Karlsson is vice president for external and United Nations affairs at the World Bank.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

issues. As a result we have received some extraordinary support from the media and, of course, from our shareholders.

Globalization and the Role of Development Economics

The Bank and the International Monetary Fund (IMF) are needed institutions in this global era, and that need will grow as globalization continues to pick up speed, making the world increasingly interdependent. But confusion is likely to grow as well. Does globalization create poverty and exclusion? Is trade bad for people, or at least for everyone except major corporations? There is not even agreement on what globalization is. We talk about it as a single phenomenon, but don't we know better? Isn't the real question how we shape globalization?

I believe that globalization refers to a number of interrelated economic, technological, social, political, and cultural changes that together have so intensified interdependence that anyone who wants to exercise control over her future must do so with others locally, nationally, and globally. But that definition is not exactly easy to communicate to a wide audience.

In fact, I think we would be better off not using the word *globalization*. We need to communicate clearly so that local, national, and global actions can be better aligned. As fear, confusion, and other emotions come into play, we need to develop ways to interact other than those seen during the protests. The people in this room will undoubtedly play a crucial role in this process—a role even more important and powerful than that played by the protestors on the streets. We need to dissect the concept of globalization to explain the phenomena that feed into it. We also need to create a better common space in which to discuss it. That is why we come together to address global issues. And we can succeed only if the debate respects other views.

There is nothing more compelling than the product of intensive research, developed with nothing else in mind than finding an enlightened approach to the objectives we want to achieve. And this conference is a golden opportunity for furthering such research.

Poverty, Policies, and Partnerships

Given the number of guests attending this conference from outside the Bank, I think it is important to address where the Bank stands on three vital issues: poverty, policies, and partnerships.

Poverty

On poverty, the World Bank has just one goal: achieving a world free of it. Over the past generation there has been a lot of progress in developing countries. Hundreds of millions of people have moved out of poverty, with a drop in the share living on less than \$1 or \$2 a day. Per capita income has doubled. The number of children in school has jumped from half to three-quarters. Infant mortality

is down 50 percent, and life expectancy is up 10 years. These achievements prove that progress is possible.

But business as usual will not eliminate poverty. Progress on primary education has stalled. Some 125 million children are not in school, and more than 500 million women are illiterate. About 1.5 billion people lack safe water, and every day 40,000 people die of preventable disease. One woman dies for every 260 live births—which I think translates into a woman a minute.

AIDS has already killed more than 20 million people, and today 34 million people are living with HIV/AIDS—23 million of them in Africa, 11 million of them orphans. Every day another 15,000 people are infected. Moreover, the disease has a dynamic in Asia and Eastern Europe that we do not yet understand. And around the world, the epidemic is having a huge effect on GNP. AIDS is more than just a health crisis—it is a development crisis, and beyond that a security crisis.

We know that trade can create wealth. Yet 80 countries containing a third of the world's population are being increasingly marginalized, and over the past 20 years developing countries' share of global trade fell from 0.8 to 0.4 percent. There are more Internet connections in Manhattan than in all of Africa. The bottom line is that 3 billion people live on less than \$2 a day, and 1.2 billion live on less than \$1 a day. We will not achieve international development goals with business as usual. And over the next 25 years the world's population will increase by 2 billion people.

So what have we learned? The Bank's *World Development Report 1990* offered a comprehensive assessment of poverty and offered three tools to reduce it: growth, education and health, and social safety nets. While these are all essential, we have overestimated the impact this strategy would have on poverty.

The 1990s taught us three important lessons about poverty reduction efforts. First, growth is crucial—but it is not enough. Second, markets free up creative capacity—but they can also be corrupt and captured. Finally, financial flows bring benefits—but the East Asian crisis revealed the vulnerability of existing systems. Alongside these lessons we have had the openness revolution (in terms of developing countries opening their political and economic systems) and the information revolution.

How should these lessons affect our work? Clearly, development requires more than just delivering money and advice, which may have been how we approached things in the past. A learning culture is needed. And we in the World Bank need to learn ourselves. To learn what development is and what delivers in development, we need to ask the poor.

So we did. In a unique collaboration, the Bank interviewed 60,000 people in 60 countries as part of participatory poverty assessments. The findings are now being published in the *Voices of the Poor* series, which is coming out in three volumes: *Can Anyone Hear Us?*, *Crying Out for Change*, and *From Many Lands*. The books are masterpieces of clear communication. If my purpose here was to engage you emotionally, I would not stop quoting from them. I advise you to read them if you have the time.

Let me run through just five conclusions of the series. First, poverty needs to be viewed in a multidimensional way. Hunger is part of everyone's understanding of poverty. But equally strong is the sense of powerlessness, voicelessness, and humili-

ation that comes with being poor. Poor people want access to basic services and infrastructure, but they know that education is their escape from poverty. Bad health is the opposite: a trigger that drags people deeper into poverty. Poor people do not just want an income or a subsidy. They know that they have to increase their assets—whether land, water, or knowledge—to get a better return on those assets.

Second, the state has been ineffective. People everywhere fear police, they hate corruption, and they trust only their own institutions. Third, nongovernmental organizations play a limited role. People rely on informal networks. Fourth, households are under deep stress. Gender relations are crucial to understanding poverty, particularly the position of men. Alcoholism is a major trap. Fifth, the social fabric is often poor people's saving grace—and it is under threat.

Taken together these findings paint a worrisome picture, and I think they partly explain why progress has not been better. *World Development Report 2000/2001* synthesizes many of these conclusions, including the three principles around which it is organized:

- Empowerment—with a pro-poor state and voice for the community.
- Security—against natural disaster, war, violence, and unforeseen changes in income and health.
- Opportunity—promoting assets and enhancing the return on them through public and private policies.

The concepts of empowerment, security, and opportunity augment our previous thinking on poverty. This thinking provides a powerful basis for moving forward with the policies needed to achieve international development goals. We can achieve those goals with more growth, trade, investment, and policies for equity and human development—and, critically, more freedom for the poor.

Policies

Thus the approach to policies must start from this broad understanding and not be based on a dogmatic approach of any kind. We must focus on results—and we know a lot about what delivers results. Such thinking underpins the Bank's Comprehensive Development Framework, which is country-led, participatory, long-term, and trying to ensure real coordination among our partners. Through it we hope to learn together, delivering development not just on the basis of money and advice.

The Bank has also tried to come closer to reality by decentralizing our operations. Today 45 country directors in the field oversee 75 percent of our lending. Lending policies have also shifted a lot since the 1980s. Take power, considered by many to be the Bank's archetypal investment activity. Twenty years ago power accounted for 21 percent of our lending. Today it is 2 percent. Twenty years ago the social sector—education, health, nutrition—accounted for 5 percent of our lending. Today it is 25 percent, or some \$5 billion a year in loans. And the total is much higher when one adds the loans we make for social protection, pension reform, labor market adjustment, social assistance, and social dimensions of adjustment.

Consider too our new ways of working with the private sector, not supplanting it but adding value. The new approach is visible in mining, oil, and gas, where the Bank does little lending but where we add value by raising environmental and other standards.

Finally, consider the surge in Bank activities involving knowledge, communication, and learning. Knowledge clusters have included InfoDev, the Global Knowledge Network, and the Global Development Network. Now we are preparing the Global Development Gateway, harnessing the Web to create a giant source of knowledge and information. This is something different, and the results are impressive. In all, this is an exciting time to be shaping the new Bank.

Partnerships

Almost none of the things I have mentioned can be delivered by the Bank on its own, especially in developing countries. The Comprehensive Development Framework, poverty reduction strategy papers, the new way we do country assistance strategies—these require partnerships at the country level and the participation of nongovernmental organizations. But at the global level something new is happening as well. Our work is becoming increasingly intertwined with the broader idea of delivering on global public policies and an emerging system of global governance.

Highlighting the many areas where the Bank is active would take too long. But in almost every field—whether promoting education for all, developing vaccines for tuberculosis or AIDS or malaria, taking a preventive approach to natural disasters, addressing the world's water crisis, or easing climate change—we are teaming up with the international community, with nongovernmental organizations, with the private sector, and with bilateral partners to deliver results in new ways.

These activities require new ways of working, and that is one of the main challenges facing the Bank. It is also crucial for you, as development economists, to help us deconstruct globalization and guide us forward.

Some of the protestors on the streets may consider the Bank to be a focus of global power. But nothing could be further from the truth. True, we deal with many actors, starting with governments. But in the modern world, power cannot rest in one locus. Whenever I have interacted with the Bank over the past 15 years, it has been a source of vibrant thought, knowledge, and learning—and that is how it should be. That is a critical global public good that the Bank can provide.

I wish you all success for the next few days of deliberation. Thank you very much.

KEYNOTE ADDRESS

Development Thinking at the Millennium

Joseph E. Stiglitz

We are entering a new millennium and closing out a century filled with ferment. As we look over the way we thought about development over just the past half century, we can see marked changes. It has become clear that development is possible but far from inevitable.

There is no single road to development. In recent years some East Asian economies have achieved sustained growth with relatively low inequality. They have not followed blindly the prescriptions of the Washington consensus: they did maintain a high level of macroeconomic stability, but at the same time their governments played a far more important role than the popular nostrums advised.

While the debate about the most effective strategies for development—and the appropriate role of the state—continues, research has helped us understand better what features of developing countries make them differ from developed countries (and they differ in many ways). We also now understand better what features of developing countries act as a hindrance to development. Today the debate is moving on to the far deeper question of how to foster change. We recognize, for instance, that what matters is not only what policies might foster faster growth, but also how the political process might produce those policy changes.

Equilibrium and *change* are thus the twin focuses of development economics. How do we describe the short-run equilibrium—the status quo, the state of being—of developing countries? And what are the forces that eventually lead to a disturbance of this equilibrium, to change?

Developing a thorough understanding of these questions is not straightforward. In my keynote addresses at the Annual World Bank Conference on Development Economics over the past three years, I have touched on several aspects of these topics, not without controversy. Economists differ markedly in their beliefs, especially on issues that touch on policy—which virtually all the questions surrounding development do. Accordingly, I have been concerned with the central question of how do we know what we know or come to believe what we come to believe.

Joseph E. Stiglitz is professor of economics at Stanford University.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

There are some policy positions on which there is almost universal agreement: for instance, governments run large, sustained budget deficits only at their peril. We can certainly agree on certain *logical* propositions, such as that free trade with perfect competition and a complete set of markets can lead to a Pareto optimum, while free trade with imperfect competition or an incomplete set of markets may result in a Pareto inferior equilibrium. But economists may differ in their judgment about the relevance of each of these propositions. And curiously, even where the weight of empirical evidence is on one side of a policy dispute, sometimes policy advisers come out in favor of the other. Or even more common, where there is a dearth of relevant empirical evidence policy advisers have come out in favor of one side with a strength unwarranted by the weight of the evidence. To understand this situation, we must pay special attention to the incentives of the policy advisers, just as we do when we study developing countries and their governments.

Equilibrium

The thinking about equilibrium in developing countries has moved a long way in the past 50 years. As in so many other areas there has been a vacillation between extremes, with current positions closer to a “middle course.”

Economic Institutions of Developing Countries—Rationality and Efficiency

Earlier models saw the behavior of people in developing countries as culturally determined in ways that made standard economic laws of little relevance. More recently, an opposite strand of research has emphasized *rationality* and *efficiency*. Evidence has shown that peasants respond to economic incentives; for instance, farmers shift production to products yielding higher prices. Earlier literature emphasized the importance of institutions that interfered with economic efficiency. The newer view saw institutions as part of the creative responses of societies to social needs—as efficient responses to market failures. A notable example is sharecropping. The earlier view saw this as an institution in which landlords not only exploited workers by taking between one- and two-thirds of their output, but in doing so also greatly attenuated incentives to produce. The newer view saw sharecropping as a rational response to the absence of market insurance and the sharecropping contract as having implicit or explicit provisions that mitigate the adverse incentive effects (see Cheung 1969).

As time went on, economists realized that many of the features of the sharecropping contract could not be explained by standard economic theories. The theories could not, for instance, convincingly explain why shares typically remained fixed even as circumstances affecting demand and supply for land and labor changed, nor could they explain why contracts were very simple.¹ Thus the concept of norms was introduced to explain institutional arrangements, bringing the analysis at least partially back to the earlier “cultural” theories.²

So, yes, incentives matter. Sharecropping provides incentives at the same time that it provides risk sharing in an economy with insurance markets that are far from complete. But there is no reason to believe that institutions themselves—the conventions that determine how individuals relate to one another—are efficient,³ or that they respond quickly or efficiently to changes in economic circumstances.⁴ Just because peasants respond to market incentives does not mean that markets are efficient. Just because an institution arises to serve a particular function does not mean that it serves that function well. (The belief that it does is sometimes referred to as the functionalist fallacy.) Markets yield Pareto efficient outcomes only under highly restrictive conditions—such as perfect information, perfect competition, and complete markets—conditions that are far from satisfied in most developing countries.⁵ Indeed, nonmarket institutions that arise to address a market limitation may actually be dysfunctional.⁶

The new approaches have not only taken the shine off the Panglossian view of rural peasants and their institutions as efficient and rational disciples of Adam Smith; they have also provided insight into these institutions, which the older neoclassical theory could not. In particular, in the past several decades there has developed a theory of rural organization explaining the structure and performance of key institutions in many developing countries, such as sharecropping contracts, rural moneylenders and, more broadly, rural credit markets, and the interlinkage of land, labor, and product markets (see, for example, Bardhan 1989; Hoff, Braverman, and Stiglitz 1993; and Stiglitz 1989b).⁷ This theory of rural organization is based on societal responses to imperfections of information and incompleteness of markets. But unlike in the earlier theory, these responses are not necessarily “efficient” or even welfare enhancing, the structure of the institutions does matter, and because the consequences of information asymmetries are affected by wealth disparities, distribution also matters.

Inequality and Poverty in Developing Countries

Anyone visiting a typical developing country could not but be struck by the huge inequalities in living standards. While a few enjoy a life of wealth and luxury, millions live subsistence lives in poverty. In some countries this disparity is a consequence of a feudal heritage; in many it is part of the colonial inheritance, a result, for instance, of European colonial masters appropriating vast amounts of land, leaving others only the residual. By a curious exercise of cognitive dissonance, when the colonial powers granted independence to their former colonies they forgot the abrogation of (implicit) property rights that they perpetrated in the first place and strongly urged that the newly independent countries respect property rights (including those of white settlers).

In the 1950s economists such as Nicholas Kaldor, Simon Kuznets, and Arthur Lewis advanced theories that inequality would lead to or was necessary for growth. Or, conversely, that growth would at least initially lead to increased inequality.⁸ These earlier economists did not discuss the impact of growth on poverty—the number and well-being of those at the very bottom of the income distribution, with incomes below some

critical level. But earlier historians and social commentators, such as Karl Polanyi (1944), had argued that in the early stages of capitalist transformation in Europe poverty had in fact increased: this happened as the social safety nets provided by communities eroded and as environmental conditions in the urban areas to which those in the rural sector were migrating degraded—before public safety nets were established (late in the 19th century in Europe, during the Great Depression in America) and before environmental regulations were imposed. Neoclassical economic theories blithely ignored issues of distribution altogether, since markets were efficient regardless of the distribution of income. Thus, since income inequality tended to remain fairly stable, increasing incomes would benefit all—including the poor.

In the past decade we have rejected most of these conclusions of “trickle-down economics.” While there is considerable evidence that growth is an important contributor to reducing poverty (see the brief discussion in Stiglitz and Squire 1998), it is also clear that some countries have been able to achieve far better outcomes for the poor than other countries with comparable incomes. Thus if a government is particularly concerned with improving the plight of the poor, it should not simply focus on growth. The experience in East Asia shows that an economy need not have huge inequality to achieve high savings rates. Indeed, the experience suggests that (at least within a range) rather than there being a tradeoff between equality and growth, the two might be complements. Egalitarian education policies, for instance, have played a pivotal role in growth (see Birdsall, Graham, and Sabot 1998; and World Bank 1993). And increased equality has led to enhanced political and social stability, thereby creating a better investment environment.

Both theory and evidence argue that distribution matters—and not just for who gets what, but for how much there is to get. Sharecropping is an inefficient economic institution that arises out of the huge inequalities in the distribution of wealth (land).⁹ The extent of credit rationing (and, more generally, the extent and consequences of capital market imperfections) also depends on distribution (see Aghion and Bolton 1997 and Piketty 1997). And a large study of sugar cooperatives in India showed that inequality in the distribution of assets affects the relative control rights of different groups of cooperative members, which in turn affects the extent to which the cooperatives depress the prices paid for inputs supplied by members and divert the resulting retained earnings (see Banerjee and others 2001).

Old concepts of “power”—discredited as Marxism grew out of fashion—are beginning to take on new meaning, especially as the links between political processes and economics move toward the center of discussion. Not long ago it was argued that it did not matter how privatization occurred in the economies in transition; as long as there were clear “owners,” Coasian arguments held, they would have an incentive to ensure that the assets were deployed in the most efficient way. While theorists long recognized the stringent conditions under which Coase’s conjecture held, the experience of the past decade has laid to rest any claims that it did not matter how privatization proceeded (Stiglitz 1999c).

Going one step further, some have also made an argument, which I have referred to as the “political Coase theorem,” that the new owners would also demand any

missing institutional infrastructure through the political process (see, for instance, Shleifer and Vishny 1998).¹⁰ Needless to say it is not oligarchs—who believe they can do better for themselves negotiating special deals behind closed doors—who demand the rule of law, but the middle classes. It was not John D. Rockefeller who strongly supported competition policy at the turn of the previous century, nor is Bill Gates doing so today.

Similarly, there is a growing recognition that the institutions set up within many societies that govern race and gender relations—including those that support discrimination in pay and access to education—have adverse effects on overall efficiency (see, for instance, North 1990). In many parts of the developing world today, institutions exist that preserve feudal power relations, institutions that interfere with overall efficiency at the same time that they increase inequality. While earlier analyses (Becker 1957) questioned how such inefficient discriminatory equilibria could be enforced, modern game theory has provided convincing and simple answers.¹¹

Features of Developing Countries That Impede Development

One reason that we want to understand the nature of the equilibrium in developing countries is that we wish to understand what may be preventing development. Why is it that so many countries seem to have “settled” on an income per capita that is so much lower than that in the more advanced industrial countries?

The older theories focusing on culture and “underdeveloped institutions” suggested that these were the barriers; only by changing culture and adopting more advanced institutions would societies bring about development—and these were daunting challenges (I return to this theme later). The neoclassical theories offered a far more positive prognosis: Institutions did not really matter. Outcomes were determined by the underlying factors of production. Countries were poor because they lacked capital. They lacked capital because of imperfections in international capital markets, imperfections begrudgingly acknowledged to exist by some of the neoclassical economists (who seemingly continued to deny the role that capital market imperfections played within countries). Thus international financial institutions had a role in facilitating the flow of capital and in overcoming the capital market imperfections that impeded the flow of capital from the capital-rich to the capital-poor countries.

Later, evidence mounted that one reason for the seeming failure of capital to flow to the capital-poor countries was that the return to capital was not high.¹² There was an obvious response: those countries lacked complementary factors such as skilled labor. This led to an emphasis on increasing education.¹³ But despite shortages of skilled labor, many developing countries have experienced alarming brain drain, suggesting that the returns to education as well as to physical capital are still higher in developed than in developing countries.

Some economists began to think that the problem was not market failure but government failure, in the form of predatory states captured by those engaging in rent seeking activities that interfered with the efficiency with which resources were

allocated. In some developing countries such interventions indeed seem to be part of the problem—but only part. Those seeking rents do not waste the volume of resources in their quest that would account for the differences in levels of production, nor does their interference seem to create a sufficiently sizable deadweight loss. Moreover, the standard arguments for rent dissipation assumed perfect competition.¹⁴ All the rents were expended in an attempt to get special treatment (such as tariff protection). But imperfections of competition characterize political markets perhaps even more than they do conventional markets. These markets are far from perfectly contestable. And in product markets it has been shown that even arbitrarily small sunk costs can enable an incumbent monopolist to maintain his monopoly power; the monopoly rents need neither be dissipated nor competed away (see, for example, Stiglitz 1987b).

These theories did not provide a convincing explanation of underdevelopment, but they did make an important contribution. Poor countries can ill afford any inefficiencies in resource allocation, and thus the elimination of market distortions would be a move in the right direction. Not surprisingly, developing countries that have moved in this direction have seen a rise in standards of living—but hardly a closing of the gap with industrial countries.

The literature also helped focus attention on issues of political economy. But there was a certain intellectual inconsistency in the stance of critics of the predatory state who argued for a minimalist role for the state. Their moral injunctions would presumably not suffice: why then would predatory states reform? When predatory states seem to reform, shouldn't political economy arguments lead to worries that the seeming reform is not really a reform, but a change in the manner of acquiring rents—and not necessarily in ways that reduce the adverse effects? (I return to this theme later.)

As we move into a new century, and as the disastrous results in most transition economies become increasingly apparent, we are coming to recognize that lack of development is often due to failures of collective action. The problem is not just predatory states but also states failing to provide the institutional infrastructure required for a market economy. Government *does* have a role—the state and the market are complements, and it is important that the state undertake its “responsibilities” and do so effectively and efficiently. The articulation of the theory of market failure, especially as applied to developing countries, allows a more precise articulation of the appropriate role of the state (see Stiglitz 1991a, 1996a, 1998b). An enhanced understanding of agency theory and of the ways in which the state differs from the market allows a better understanding of some of the mechanisms by which the efficiency and effectiveness of the state can be improved (see Stiglitz 1989a, 1991b).¹⁵

The state, besides ensuring that it performs its own functions well, also has a role to play in helping to alter other institutions in a country. Once it is recognized that the institutional arrangements in a society might not be efficiency enhancing, but rather power and wealth preserving, a new rationale for collective action arises—not just to correct market failures but to alter nonmarket institutional arrangements

that impede efficiency and increase inequality. These institutional arrangements are sometimes, but not always, creatures of political processes, for instance, at the local and village levels. But to alter arrangements at one level might require interventions arising from another level.

The Role of the State in Altering Equilibrium Conditions

So far we have followed the development debate as it has cycled from institutions and culture explaining everything, to their being irrelevant, to their being explicable themselves and efficiency enhancing, to their being partially inexplicable and possibly even dysfunctional. We have followed the debate from inequality being unimportant, to inequality promoting growth, to inequality having a direct adverse effect on growth—and to a recognition that institutions, rather than promoting efficiency by filling in the gaps from market failures, may even impede efficiency as they seek to preserve existing inequalities and power relationships. Conversely, there has been a shift from the view that growth inevitably increases inequality, to the view that growth inevitably improves the plight of the poor, and, finally, to the view that while growth normally improves the plight of the poor, that outcome is not inevitable and some growth strategies are more pro-poor than others.

We have seen too how the explanation of why developing countries are poorer than developed ones has shifted—from capital, to human capital, to predatory states. But while each of these plays a role, there is a sense that (at least as conventionally articulated) the differences are too great to be explained just by these factors. Some other factor or factors are also at play.

If we write the aggregate production function in the standard way,

$$Q = F(A, K, L, H)$$

where K is capital, L is labor, and H is human capital, there is a “factor” A that differs between less and more developed countries. Let us think of it as a vector of dimensions that characterize countries as being more or less developed. In the rest of this section I illustrate some of the key components of the “factor” A .

INFORMATION. The theories discussed so far suggest that imperfections in information help lead to an equilibrium in developing countries in which resources are less efficiently allocated than in industrial economies. The imperfect information (agency) theories and the “market failures” they bring to light suggest four immediate implications for the role of government (see World Bank 1999):

- Information flows are weaker in developing countries. Governments should take action to improve information flows, which would in turn affect prevalent institutional arrangements.
- Agency problems, such as those associated with sharecropping, reduce production and consequently the welfare of all participants. Governments should take actions to reduce agency problems and the consequences of

information asymmetries more generally; one such action would be redistributing land so that farmers own their own land.

- Weak capacity for gathering and processing information in developing countries prevents the establishment of many necessary market arrangements enjoyed in developed countries. Governments should help create institutions that find more efficient ways around information asymmetries, such as microcredit schemes with peer monitoring. (For an overview of microcredit see, for instance, Morduch 1999; for an early theoretical analysis of government strategies see Stiglitz 1990.)
- In general, information is imperfect and markets consequently work imperfectly. Thus in the design of all policies governments should take into account that markets cannot be assumed to work as they would under the standard “demand and supply” framework.

KNOWLEDGE. Another dimension of *A* is knowledge: developed countries differ from developing countries in their knowledge, including that of production processes. Once again government has a role in promoting the generation, acquisition, and use of knowledge.¹⁶

- Knowledge is like a public good, or at least an impure public good (for an early discussion of this issue see Stiglitz 1987a). Thus there is likely to be underinvestment in the production and dissemination of knowledge, giving rise to an important public role in these activities.
- The transfer and adaptation of new technologies, central concerns of developing countries, are no less a public good, and no less important, than the original production of knowledge. Governments can take a cue from East Asian economies and invest in a highly educated, technologically literate labor force and in the transfer of technology (World Bank 1993, Stiglitz 1996b).

The World Bank has increasingly seen closing the knowledge gap as one of its central roles. The president of the World Bank, Jim Wolfensohn, has often spoken of it as a “knowledge bank” (see Wolfensohn 1996). Knowledge, including knowledge about development, is a global public good: the benefits of that knowledge can be of value to everyone. It is therefore appropriate that knowledge be provided by an international public institution, such as the World Bank.¹⁷

SOCIAL AND ORGANIZATIONAL CAPITAL. A third dimension of *A* is a country’s social and organizational capital. Production processes take inputs and translate them into outputs. That process typically involves more than one individual, requiring individuals to relate to one another. How they relate affects the efficiency with which inputs are transformed into outputs. And how they relate depends on their information, on the organizations in which they work, and on the norms and institutions that govern their “public” behavior (that is, their behavior in dealing with others). In modern societies we typically take these forms of “capital” for granted, though we often appreciate people whose word is their honor (so that no contract is

required). But in the transition economies and in many developing countries we have seen what happens when social capital deteriorates.¹⁸ In both the Russian Federation and African countries, for instance, electric lines fall prey to copper thieves.

A large literature emphasizes the importance of contract enforcement; without such enforcement intertemporal trades cannot be made and credit markets cannot work. Government has a role in ensuring a strong and consistent judicial system, but most contracts are not enforced through the courts. Reliance is placed on self-enforcing incentives for compliance, associated, for instance, with repeated games and with norms, the breaking of which results in adverse social and economic consequences. Modern analyses of organizations have emphasized that the “good” behavior of organization members often seems difficult to explain simply by looking at incentives—even long-run incentives associated with job promotion and reputation (for classic analyses see Marshall 1897). Organizations seek to change the “preferences” of members, to have them identify with the organization (see Simon 1991).

The level and nature of social and organizational capital certainly differ across countries, and they can change during development and transition—often in adverse ways.¹⁹ Some of these effects are an inevitable consequence of development—but even then it is important for policymakers to be aware of them and possibly to take offsetting actions. (In a few countries, such as Uzbekistan, the government has explicitly recognized the importance of social capital and tried to strengthen it.) In some cases governments have some choice over policies (such as having extremely high interest rates) that may produce an additional adverse effect.

Poverty Traps

So far we have seen how norms, the distribution of wealth, and institutions (including political institutions) affect equilibrium in developing countries.²⁰ There are a host of factors, besides differences in capital per worker, that explain differences in output per worker.

Another line of analysis argues that there can be more than a single equilibrium. Which equilibrium the economy “chooses” may be determined by history—and affected by government action.²¹

A recent strand of analytical research has focused on the observation that the major determinant of the environment of each actor (firm or household) in any economy is the behavior of other actors, and that there may be multiple equilibria.²² Charles Darwin ([1859] 1993) observed something similar in his visit to the Galapagos Islands, a set of islands that have similar physical characteristics but developed markedly different flora and fauna.

This view that development is not deterministic is in direct contrast to the standard neoclassical approach. In that approach the economy at each date is determined by preferences, technologies, and resources. And the story of its evolution is equally simple: today’s preferences, technologies, and resources determine those of

tomorrow.²³ Policy has an accordingly limited scope: it can only help facilitate the transfer of resources (or technology) or increase the rate of their accumulation (for instance, through actions that increase the savings rate).

By contrast, in the multiple equilibria models another set of actions is possible: the government can impose policies that help move the economy from one equilibrium to another. For instance, if families worry about unemployment, they may send more than one worker into the labor force. Thus an increase in the unemployment rate leads to an increase in the supply of labor. If there is a minimum wage (or an efficiency wage) at a level above that at which demand for labor equals supply, there may be one equilibrium with a low unemployment rate, so that each family sends few workers into the labor force. There may be another equilibrium with a high unemployment rate, so that each family sends many workers into the labor force. But if the government were to provide a guaranteed income to families, each family would need to send out fewer workers, and the only equilibrium to emerge would be the low unemployment equilibrium (see Basu, Genicot, and Stiglitz 1999).

The theory behind affirmative action is similarly that a government intervention can help eliminate a “bad” equilibrium, with statistical discrimination. Multiple education equilibria can exist: If relatively few people choose to get educated (say, at a high level), the average productivity of those without education will be relatively high and the returns to an education low, so that only those for whom the cost of acquiring (a high level of) education is low choose to do so. But if many choose to get educated (again at a high level), the average productivity of those without education will be relatively low and the returns to an education high, so that it will pay a large fraction to obtain (a high level of) education. If there are two racial or ethnic groups in a society, one may wind up in the low equilibrium trap. This is possible particularly if the members of that group had at one time been precluded from large segments of the labor market, so that for them the return to education was low. With enforcement of nondiscrimination or affirmative action policies, the minority comes to be treated like the majority, in which case the minority’s low-level equilibrium is eliminated (this model is developed in greater detail in Stiglitz 1974b).

Dynamics of Change

We now have a broader understanding of why developing countries may have such a low level of income. The reasons go beyond a lack of capital. As we have just seen, there may be a “poverty trap.” The harder question is what drives change—what enables it to occur in ways that move an economy from underdevelopment to development.

Development as Transformation

First I should say something about what we mean by “development,” which follows from much of what has been said in the previous section. Developing countries, in the neoclassical vision, were just like developed countries except that

they were poorer; they had fewer resources—in particular, less capital and human capital. So there was little to development economics: it was simply a question of what could be done to increase the pace and efficiency of the accumulation of capital.

From this perspective, Hollis Chenery and Anne Krueger, two of the earlier chief economists of the World Bank with different viewpoints, had much in common. Both the planning approach associated with Chenery and the approach arguing for liberalization associated with Krueger saw the essential problem of development as a narrow one of resource allocation—increasing the efficiency of resource allocation and allocating more resources to investment. One, noting the pervasive market failures in developing countries, looked to planning as a way of increasing the efficiency of resource allocation and increasing investment. The other, noting the pervasive government failures in developing countries, looked to stripping away the role of the government, hoping that the market, left to itself, would both allocate resources more efficiently and allocate more resources to investment.

More recently there has been a return to older doctrines that saw more to the development process.²⁴ Development is now seen as a transformation of society, a move from old ways of thinking, and old forms of social and economic organization, to new ones (see Stiglitz 1998a, b, c).

The new view argues that development and the developmental transformation involve a change in the way people think and the way societies function—a change in norms, expectations, and institutions. These changes are reflected in the *A* of the production function: more developed countries are able to transform a given set of inputs into greater output. But the changes have even more profound effects on the dynamics of change, for development involves not just the acceptance of change but its promotion and, indeed, its routinization. Traditional societies typically accept matters as they find them; the developmental transformation entails questioning existing arrangements and continually seeking alternative and more efficient ones—including better processes of change, reflected in what is typically referred to as the scientific method. All societies involve a mix of the new and the old, of scientific and prescientific ways of thinking. In developing countries, however, traditional modes of thinking, and traditional institutions, dominate.

This more expansive vision does not bring with it a corresponding set of obvious strategies for promoting change. But it does point to some areas that may have received insufficient attention, and give new reasons for certain development policies. Education is important not just because it increases “human capital,” but because it also increases the acceptance of change—it introduces individuals to the scientific method, to ways of thinking that are markedly different from traditional ways. Trade is important not just because it makes it possible to purchase some goods at a lower price (or makes some goods available that otherwise would not have been). Trade is also important because it brings a country into closer contact with others, and these contacts bring a change in ways of thinking. In this respect promoting exports (making countries understand the nature of the international

marketplace, the importance of standards, and the like) may be far more important than liberalizing imports.

Reform

Perhaps the hardest question facing policy economists interested in development relates to reform: how to change the institutions and policies that govern an economy. What disturbs an equilibrium and causes the economy to move from an underdevelopment equilibrium to a developmental path characterized by greater wealth creation and rising standards of living?

Even if we think of the government as an “outside actor” (rather than as part of the system), we must contend with the difficult problem of how the government can change ways of thinking and institutional arrangements. But matters become even murkier if we think of the government as endogenous.

The nature of the problem is posed most clearly by some of the recent debates on privatization. Governments in many countries have run enterprises that produce conventional goods, often at low efficiency. One reason cited for the poor performance is corruption, government officials’ continual rake-off of the profits of public enterprises. The model for change that seemed to underlie the strategies of the international financial institutions was as follows: A mission (made up of something akin to missionaries) would go to the country and explain the virtues of privatization—in particular, how privatization would increase efficiency and stem the corruption that was bleeding the economy. The government officials (the perpetrators and beneficiaries of the corruption) would suddenly see the light, cry out “hallelujah,” and run to the parliament to pass a privatization law. The feelings of virtue that would overcome the government officials would more than compensate for their loss of income.

In reality, of course, privatization allows corrupt government officials to steal not just a fraction of today’s rents (profits), but a fraction of the present discounted value of all future rents. If government leaders are corrupt, inferences about policy reforms should be made with care. The reforms may be endorsed because they enhance the opportunity for corruption, not because they promote the overall efficiency of the economy. All this, one might say, is elementary economics: incentives matter, and we should look at the incentives of those in power. While elementary, this principle seems all too often to have been ignored. And in doing so, the international financial institutions may in some instances have advocated policies that had the effect of aiding and abetting corruption rather than stemming it.

If this is so, it raises a difficult issue: when and how will meaningful reform occur? Here I want to put forward a simple way of looking at this issue that focuses on *incentives* and *ideas* as they affect both individuals and the coalitions that form among individuals and groups, and examines whether they work to promote or resist change. I also attempt to identify the disturbances to the initial equilibrium.

Several examples will help illustrate. In the example I just gave, the disturbance to the initial equilibrium was an *idea*, the notion that privatized firms are more effi-

cient than government-run enterprises. Unfortunately, this idea was closer to being part of an *ideology*, a primitive (nonscientific) belief system, not well rooted in evidence or theory. And the *incentive* to implement that idea was the ability of the politician to appropriate even more rents. Thus the transition economies that were encouraged to privatize and to liberalize their capital accounts before establishing good corporate governance and other elements of the institutional infrastructure required for a market economy truly did change. But the result was that politicians (acting according to their incentives) privatized by selling state assets to those most willing to pay them bribes, and those buyers (acting according to their incentives) promptly stripped the assets from their new firms and took a substantial part of the value out of the country (see Black, Kraakman, and Tarassova 2000).

The provision of public education is an example of a reform driven by the *incentives* of ruling elites that eventually led to a reduction of their power. Here the disturbance was the development of new technologies that required more trained labor. With a sufficiently high discount rate, supporting more mass education would pay off for existing elites even though it would mean a loss of power in the future. (For a more extensive discussion of this example see Bourguignon and Verdier 2000.)

During the East Asian crisis the *idea* of transparency surfaced, the lack of which was cited by some as one of the main causes of the crisis. Several contrary observations, however, led to a widespread suspicion that transparency was not at the root of the problem.²⁵ (Perhaps those who raised concerns about transparency were themselves trying to hide something!)²⁶ But once the idea of transparency was let loose, it took on a life of its own. Thailand, for instance, included in its new constitution citizens' basic right to know. The need for a comprehensive disclosure regime became recognized, a regime that would embrace hedge funds and offshore banking centers as well as the banks in the developing world. But the arguments for transparency seemingly came full circle, and now some of those who had spoken so loudly in its favor (for instance, in the U.S. Treasury) expressed some misgivings—going so far as to suggest that excessive disclosure requirements might be counterproductive, since they would reduce the incentives for gathering information!

The authors of America's declaration of independence, who wrote "All men are created equal," probably did not fully grasp the implications of this fundamental *idea*. But it was an idea with enormous power, one that was eventually used to attack slavery and discrimination in all its manifestations.

Evolutionary Change

In the previous section I emphasized the distinction between change and equilibrium. The analysis of developing countries entails examining why the equilibrium in many of these countries appears to differ so markedly from that in more developed countries and what the forces for change are. In fact, much of traditional analysis has focused on equilibrium dynamics, that is, change that is itself part of an equilibrium process. By contrast, I have contended that the focus should be on a developmental

transformation that is more fundamental—that involves changes in individuals and societies, in their ways of thinking, and in institutional arrangements—and that such changes cannot be described well by models of equilibrium dynamics. The difficulties of analysis have been highlighted in my attempt to describe reform processes, the deliberate attempt by government or outsiders to intervene in ways that lead to these more fundamental transformations.

A distinction is sometimes made between revolutionary and evolutionary processes of change. The first entails rapid changes, affecting the foundations of society and its institutions; the second entails more gradual changes, affecting only a few parts of the system at a time—though over time the cumulative effects of such changes can be revolutionary. From this perspective the American revolution was not really a revolution, for social and economic relations remained largely intact; indeed, it can be viewed more as an attempt to prevent a reversion to an earlier social order than as an effort to force a change to a new order. The French, Russian, and Chinese revolutions, at least in their conceptions, were far more revolutionary; they purported to fundamentally change social relations. Each encountered problems—excesses followed by changes that resulted in a social order different both from what had prevailed before the revolution and from the conception of those who had initiated the revolution. Recognition of the problems encountered by revolution has renewed the focus on evolutionary processes of change.

The analysis of evolutionary processes begins with the observation that societies are always bombarded by shocks that give rise to change. Some forces for change arise internally, others from the outside.

INTERNAL SOURCES OF CHANGE. Among internal sources of change are the processes of research and development in modern society; these generate changes in technology that may in turn lead to profound changes in society. The Internet revolution portends to be every bit as important as the industrial and scientific revolutions that preceded it, fundamentally changing the ways in which individuals and businesses interact with one another.

Ideas often take on a life of their own. They evolve over time, an evolution that is both the consequence and the cause of changes in society. Ideas, simple and complex, gradually find applications in one area after another, evolving in the process. The full implications of such ideas as assembly lines, just-in-time production, and replaceable parts can take years, even decades, to play out. As they do, change affects one part of society, then another—in a never ending progression—until virtually every part has been transformed.

EXTERNAL SOURCES OF CHANGE. Some of the most profound changes in societies have stemmed from events that are at least partially exogenous. The encounters in the 19th century between, for instance, Asian and European societies led to major changes, particularly in the Asian societies as they confronted new ideas and new circumstances (including new technologies). The recent global financial crisis has left in its wake a firmer commitment in many countries to democratization, a firmer

opposition to corruption, but also a broader suspicion of ideology, as the adverse effects of premature capital market liberalization—so strongly supported by market ideologists—have gradually sunk in.

CHANGES CANNOT BE FORCED. Among the sources of change, as I have said, are changes in knowledge and information or, more broadly, in beliefs and perceptions. But such changes cannot be forced. Encounters across cultures have such impacts in part because they indirectly change people's beliefs about what is possible. Outside advisers can thus have the most profound and lasting effects not through conditions attached to loans and aid (which are unlikely to be sustainable through the vicissitudes of changing political currents) but through analysis—by exposing the alternatives and the risks and consequences of each. Arguments based on ideology are most successful in swaying only those who are already converted, not those who have yet to see the light.

Science, Ideology, and the Developmental Transformation

The developmental transformation has already changed people's ways of thinking in many parts of the world. The scientific method has been embraced, and there is little enthusiasm for abandoning one religion only to adopt a new one of market fundamentalism. As I said in my address at the 1998 Annual World Bank Conference on Development Economics, an explicit part of the scientific method is the recognition of uncertainty (Stiglitz 1999b). By failing to convey to the developing countries the range of views and the sense of scientific uncertainty, we do them and ourselves a disservice: not only do we set back democratic processes and the developmental transformation, we undermine our own credibility. There is a marked contrast between the seeming confidence and certainty with which the international financial institutions impose certain reforms and the critical scientific method, which fosters the

willingness to hold belief in suspense; ability to doubt until evidence is obtained; willingness to go where evidence points instead of putting first a personally preferred conclusion; [and] ability to hold ideas in solution and use them as hypotheses to be tested instead of as dogmas to be asserted . . .
(Dewey 1939, p. 145)

This part of the scientific attitude is translated into the policy domain through such suggestions as multiple advocacy (Haas 1990, p. 210) and double visioning (Schön 1983, p. 281).

Why do the international financial institutions so often ignore these precepts, presenting their advice as if the policies they recommend were Pareto dominant policies and without sufficient warning of the associated risks and the limitations of the scientific evidence in their support? This question is particularly relevant because of the irony: the prevalence of prescientific, ideological ways of thinking in institutions

supposedly committed to bringing about the developmental transformation, an essential aspect of which is scientific thinking.

Only recently have such institutions become subject to the same analytic scrutiny to which other organizations and institutions (including governments) have been subjected. Such analyses begin with the presumption that it is not wanton perversity that leads to these behavioral patterns in large international organizations. There are quite human impulses that push for conformity and rigidity.

As economists, we believe strongly in the importance of incentives. In my 1998 address I suggested that we economists and the institutions for which we work might not rise above this general principle. The incentive structures we face—and, more particularly, the interests we and the institutions we work for serve—might play more than a small role in the advice we give and in the views of the world that we hold. Methodologies that support those views are encouraged; studies that contradict them are given more intensive scrutiny. Is it any accident that the cross-sectional studies showing that trade liberalization is positively associated with growth were widely disseminated, while those showing that capital market liberalization is not so associated were not (see Rodrik 1998)?²⁷ The reason I touch on this epistemological issue here is to encourage a healthy skepticism.

To be sure, we must guard against the opposite danger, of pretending that there are no general truths, no lessons to be learned, from economic science. Countries must make decisions, and they want to base those decisions on the best available evidence. They might like it to be the case that economic science pointed the way to particular decisions. In some areas there is indeed a broad consensus among economists. But in other areas economists may disagree even where the weight of empirical evidence is on one side. For instance, the evidence shows that capital market liberalization does not enhance economic growth and does increase risk. Yet many policy economists, especially at the International Monetary Fund (IMF), advocated capital market liberalization, though now (too late for many who have suffered from premature liberalization) they seem to openly recognize the risks.

A natural question is, why, in the absence of evidence of its benefits and with strong evidence of its risks, did the IMF still advocate capital market liberalization with such vehemence? The easy answer would be ideology—IMF economists' strong belief in market fundamentalism led them to ignore the dangers of this policy prescription. While this answer is probably true, an economist naturally pushes the analysis further: Were there incentives at play? Were some interests being served? Framed this way, the question answers itself once one looks at the governance structure of the IMF. The organization is dominated (in terms of voting power) by the advanced industrial countries, each represented by officials from its central bank and finance ministry, with close links to the country's financial markets. And the United States, the single country with effective veto power, stands to gain much from increased access to financial services markets.

The advocacy of capital market liberalization is an extreme case, both because of the clear evidence against the policy and because of the clear interests of those dominating the IMF. In other areas, relevant empirical evidence or well-developed the-

ory in support of a policy choice may be lacking, but all too often policy advisers nevertheless come out strongly in favor of one side or the other. For instance, problems in public pension programs made it clear that reforms were desirable. But it does not follow from the fact that some public pension programs fail that all of them must fail. And it is a fallacy that government failure requires a market solution. Even in developed countries transactions costs, potential abuses of uninformed investors by unscrupulous firms, and high market volatility (combined with the unavailability of insurance against such risks as inflation) may make privatization unattractive (see, for example, Orszag and Stiglitz 2000).

All these problems are far more severe in developing countries. Thus policymakers often face the difficult choice between strengthening a publicly managed pension program (possibly by using private contracting) and privatizing it while trying to establish a regulatory regime that prevents abuses and insurance schemes that provide a modicum of economic security to the aged. Outsiders from international agencies have tended to strongly recommend the second choice, even though there is little basis for this stance. Why? Again, ideology provides only a partially satisfying answer; the role of interests cannot be ignored.

Resistance to Fundamental Change

Thus all too often advice has been presented that goes well beyond the evidence in its support. But there is a concern too about the way in which advice is given—and, more broadly, about how policy decisions are made. Why is only a single view presented? Why does so much go on in secret, without broad consultation? Besides the human proclivity for establishing and defending authority, there is another force that shuts down inquiry, and that is the recognition by those in power that new ideas can be a force for change—and for changes that undermine that power.

Because knowledge and information can be such a powerful force for change—and for changes in direction that cannot always be fully anticipated—those in power have a strong penchant for conducting public affairs in secret. They want to control, as much as possible, the process of change—to make sure that it does not take a turn not in their interests. The more open and public the decisionmaking process, the greater the possibility of such unwelcome turns in events.

Much of the change in society is effected through political processes. Whether changes occur, and what changes occur, depend in no small measure on what coalitions form. But the dynamics of coalition formation are complex, affected by incentives, ideas, and beliefs and expectations about, for instance, what changes are feasible now and in the future.

The dynamics of coalition formation are important both in understanding resistance to change and in thinking through the consequences of alternative strategies for sequencing change. And they are important both for those wishing to prevent change and for those wishing to promote it. For instance, elsewhere I have suggested that often some groups recognize that a certain reform would result in a political dynamic posing the risk that they would be worse off in the long run even if they

would be better off in the short run (Stiglitz 1998d).²⁸ Having recognized this dynamic, these groups often block the reform.

Political Processes and the Promotion of Change

To see how awareness of the dynamics of political processes affects the design of reform strategies, consider a government privatizing a monopoly, with the eventual goal of creating a competitive private market. Should the government wait to privatize until after a competition law has been passed? One view would be that delay would only extend the deadweight losses from inefficient public ownership. But once political dynamics are taken into account, delay might appear more desirable—for without delay, the ultimate objective might never be attained: it might be easier to adopt a competition policy before a noncompetitive privatization has occurred than after. A privatization that results in a monopoly firm has created a vested interest against competition. But behind the veil of ignorance—before anyone knows who will become the monopolist—all can agree that it is better to have more competition.

I have argued that sustainable development, and sustainable reform, are based on changes in ideas, interests, and coalitions. Let me repeat that *such changes cannot be forced*. Changes in ways of thinking often take time. That is why the approach to reform based on conditionality has largely failed (see World Bank 1998). That is why the Bolshevik approach to changing society—forced changes from a revolutionary vanguard—has failed time and time again. The shock therapy approach to reform was no more successful than the Cultural Revolution and the Bolshevik Revolution.²⁹

Conclusion

To return to the refrain at the beginning of this essay: We now recognize that development is possible but far from inevitable. We have learned that there is more to development than more rapid accumulation of capital, though without that development is unlikely to occur. And we have learned that there is more to development than more efficient allocation of resources, though poor countries can ill afford to waste any resources.

The most successful developing countries have not followed the neoliberal doctrines encapsulated in the Washington consensus. To be sure, growth is difficult without macroeconomic stabilization. But Turkey has shown that it is possible to have sustained (though precarious) growth even with rather high inflation, and econometric studies have shown that reducing inflation below a critical threshold yields few if any discernible benefits, though disinflation may have large costs. The recent crisis in East Asia has reminded us—if we needed reminding—that economic instability may arise from a multitude of sources other than bad macroeconomic policies. Indeed, it is increasingly recognized that some of the same policies that the international financial institutions pushed in the name of promoting growth also increased economic volatility (see Easterly, Islam, and Stiglitz in this volume).

All too often there has been a confusion between means and ends. Liberalization and privatization have been pursued as ends in themselves rather than as means to more rapid and equitable growth or greater economic stability—and all too often they have been pursued in situations where the consequences worked against those more basic objectives.

Experience in developed countries should have provided a clear warning that there is far more to development than privatization, liberalization, and stabilization. The South of Italy has lagged far behind the North—one of the most dynamic regions in the world—despite there being no trade barriers with the rest of Italy (or the rest of the European Union). The economic regime of the South—in terms of liberalization, privatization, and stabilization—is essentially the same as that in the North, yet its performance is markedly different. Something else matters, and matters a lot.

We now recognize too that long-term sustainable growth requires the development of a consensus behind reform policies. *Reforms cannot be imposed from the outside*—part of the reason for the widespread failure of conditionality. Meaningful democratic processes that involve participation and voice, combined with policies that promote equity, can enhance consensus building and a sense of inclusion and even create the social capital that is increasingly recognized as a key to long-term success. The developmental transformation entails more than solving technical economic problems so that an economy can increase efficiency and resource mobilization—as important as this is. Thus in our thinking about development, we have gone beyond projects—recognizing that even good projects will make a difference only if they are scaled up—and beyond policies—from good macroeconomic policies to dynamic programs for enterprise creation. We have begun to focus on institutions, such as those that promote competition and good governance in the public and private sectors. Contrary to the neoclassical model, institutions do matter.³⁰

One important way in which the new theories differ from the old is in stressing that not only institutions matter, but also the distribution of income. Thus reducing poverty and inequality assumes an importance not only as an end in itself, but also as a means of achieving stronger economic performance. The World Bank's *World Development Report 2000/2001: Attacking Poverty* has emphasized the multidimensional nature of poverty—the poor face not only a chronic shortage of income but also a sense of voicelessness and powerlessness and a high level of economic insecurity (World Bank 2000). The discussion continues about the extent to which policies pursued in recent years have promoted economic growth, but there is little debate that in at least some instances the policies and the way in which they have been imposed have increased economic insecurity and the sense of powerlessness.

As we enter the 21st century, we thus approach development with a more comprehensive framework, an awareness of broader objectives and more instruments, a greater sensitivity to the complexity of the development process, a greater sense of humility in the face of the tasks ahead—but also a greater sense of optimism about

what the future might bring.³¹ We know that development is possible. The challenge is to foster it in ways that benefit the poor, strengthen democratic processes, heighten the overall sense of well-being, and widen economic and political freedom.

Notes

1. In general, optimal contracts were highly nonlinear (see Peyton 1996). To be sure, economic theorists did provide partial explanations of this phenomenon; for instance, the possibility of arbitrage (and the impossibility of monitoring such arbitrage activities) meant that it might be impossible to implement nonlinear contracts, or contracts in one region with shares differing markedly from those in a neighboring region.

2. See, for instance, Platteau (1994). One manifestation of this concern with norms is reflected in the recognition of the importance of social capital (see the section on social and organizational capital).

3. Thus neoclassical theory argued that institutions either did not matter—economic outcomes were always efficient, and one had to see through the veil of institutions to the underlying economic forces (as Cheung had argued in the case of sharecropping) or mattered only to enhance economic efficiency. Stiglitz (1974a) showed that, with imperfect information, both propositions are wrong; sharecropping did attenuate incentives and the sharecropping equilibrium was not, in general, constrained Pareto efficient (Greenwald and Stiglitz 1986).

4. There have been some attempts to explain the failure of adoption on the basis of theories of asymmetric information. See, for instance, in the context of the evolution of contractual provisions, Stiglitz (1992).

5. Markets are not even constrained Pareto efficient, that is, taking into account the costs of establishing and running markets and of obtaining information. See Greenwald and Stiglitz (1986).

6. Arnott and Stiglitz (1991) show for the insurance market that the nonmarket institutions, which arise from the partial insurance that arises from concerns about moral hazard attached to market insurance, may crowd out market institutions, so that the equilibrium level of market insurance—and, more important, the level of expected utility—may be lower.

7. Other strands of literature have provided insights (often in terms of the responses to risk) into demography (the determinants of family size)—see, for instance, Rosenszweig (1988)—and other aspects of household behavior.

8. Kaldor (1956) was one of several authors to put forth the idea that wage earners and profit earners had different propensities to save. According to these theories, only by increasing the income share of the rich (or of capitalists) could one increase the aggregate savings rate. Kuznets (1955) described how in an economy that drew more and more people from (low-income) agriculture into (higher-income) industry, inequality would first increase, then decrease. Lewis (1954) set out the dual economy theory of development, which stressed the importance of savings: increased inequality, increased aggregate savings, and thus promoted growth.

9. This is an example of a more general set of propositions in agency theory, which showed that the clean separation of distribution and efficiency issues that characterized neoclassical theory was no longer true when there was imperfect information. See Hoff (1994) and Stiglitz (1994).

10. Though to be fair to Coase, he was much more aware of the limitations of this “theorem” than were most of those who invoked it, and it is unlikely that he would have gone so far as to subscribe to the tenets of the “political Coase theorem” (see Coase 1992).

11. Modern game theory has reinforced earlier, more intuitive arguments on this question. See, for example, Abreu (1988) and Akerlof (1985).

12. It certainly was not as high as the shortage of capital would suggest it should be, using conventionally estimated production functions. See, for example, Lucas (1988) and Stiglitz (1988).

13. Again, capital market imperfections played a key role in explaining underinvestment in education: the poor especially lacked access to capital to enable them to invest in education, despite the huge return to such investment. The inherent problems in correcting this market failure meant that in most countries there was a presumption for a large role for the state.

14. See Krueger's (1974) seminal article on government rent seeking and Bigsten and Moene (1997) for an example of rent dissipation.

15. For practical "manuals" see the discussions of the "reinventing government" initiatives in the United States, for instance, Osborne and Gaebler (1992). For an application to developing countries see Stiglitz (1998b).

16. Standard neoclassical theory assumed that knowledge was given and fixed—a particularly inappropriate set of assumptions for developing countries. It was also assumed that knowledge moves quickly across porous boundaries. Yet even within countries there seem to be large differences in productivity between "best practice" firms and average firms, differences that cannot be easily explained in terms of the standard factors of production.

17. For a discussion of the concept of international public goods and the role of the international financial institutions in the provision of international public goods, including knowledge, see Stiglitz (1995b, 1999a, 2000b).

18. For a discussion of social capital in the context of a comparison of development in China and Russia see Hussein, Stern, and Stiglitz (2000). For an excellent recent overview of social capital see the collection of papers in Dasgupta and Serageldin (2000). Also see the references at www.worldbank.org/poverty/scapital/library/keyread1.htm.

19. Some of these changes can be understood in conventional economic terms: high interest rates reduce the value of reputational capital; periods of rapid change, with high probabilities of organizational death, reduce the incentives to maintain organizational capital (see Stiglitz 2000a). Other changes are beyond the realm of standard economic analysis.

20. This section is based on Hoff and Stiglitz (2001) and the papers cited there.

21. Gunnar Myrdal (1957, 1968) has probably done the most to popularize the intuitive notion of the vicious circle in the economic development literature, particularly to account for persistent and increasing national and international inequality. Positive feedback means that some deviation or tendency is reinforced so that the process "snowballs" until it meets some countervailing force. For developing countries, processes of "catching up" can be accelerated and processes of "falling behind" aggravated (at least in relative terms). A virtuous circle and a vicious circle operate to drive the dynamics of divergence in opposite directions. The vicious circle leads to a low-level equilibrium and the virtuous circle to a high-level equilibrium. To get out of the trap of the low-level equilibrium, a comprehensive development effort of coordinated action is needed so that the virtuous dynamics take hold and drive to the high-level equilibrium. Models of multiple equilibria associated with "coordination failure" are associated with Rosenstein-Rodan (1943) and more recently, Murphy, Shleifer, and Vishny (1989).

22. If almost everyone else in society is (or acts) bureaucratic, it is more likely that it pays me to be (or act) bureaucratic, and it is more likely that those with bureaucratic mentalities will prosper and multiply. There is thus a bureaucratic equilibrium, but there may also be an innovative equilibrium, in which most individuals are "innovative" (see Stiglitz 1995a). In the innovative culture bureaucratic people do not survive, and conversely in the bureaucratic environment. There are thus multiple equilibria (in this case, multiple equilibria cultures). Robert Putnam (1993, p. 177) wrote about the self-reinforcing cultures of northern and southern Italy:

Stocks of social capital, such as trust, norms, and networks, tend to be self-reinforcing and cumulative. Virtuous circles result in social equilibrium with high levels of cooperation, trust, reciprocity, civic engagement, and collective well-being. These traits define the civic community. Conversely, the absence of these traits in the *uncivic*

community is also self-reinforcing. Defection, distrust, shirking, exploitation, isolation, disorder, and stagnation intensify one another in a suffocating miasma of vicious circles. This argument suggests that there may be at least *two* broad equilibria toward which all societies that face problems of collective action (that is, *all* societies) tend to evolve and which, once attained, tend to be self-reinforcing.

23. Note that nothing in neoclassical theory itself ensures a unique equilibrium, but the simple aggregative models typically had structures that guaranteed that outcome.

24. For instance, that of Paul Rosenstein-Rodan, the World Bank's first chief economist. See, for instance, Rosenstein-Rodan (1943).

25. Note that the last set of crises occurred in the Scandinavian countries, countries with the seemingly highest level of transparency. This certainly suggests that transparency itself does not inoculate against crises. Also note that transparency had been increasing in many of the East Asian countries, and that many other countries not experiencing a crisis had far less transparency. Finally, note that most of the relevant information (for instance, Thailand's persistent trade deficit and Korea's high level of corporate indebtedness) was already widely known, and that economic theories even suggested that more information might be associated with increased market volatility. For a review of these arguments see Furman and Stiglitz (1998).

26. Indeed, there were clear incentives for pushing the transparency argument: the investment firms that had pushed their clients into investing in these countries wanted to shift blame (it was not their faulty investment advice that was to blame; what were they to do, given the lack of transparency in the countries?). To be sure, there was a certain hollowness in this argument, since these investment advisers should have, at the very least, been aware of the lack of transparency. And to a large extent the only reason that they could make higher than normal risk-adjusted returns on these investments was the investors' purported informational advantage. Some industrial country governments, especially the U.S. government, had an interest not only in defending these firms but also in shifting blame: they did not want the spotlight to turn on them and reveal that the underlying problem was the excessively rapid financial and capital market liberalization that they had pushed on these countries.

27. Similarly, empirical studies, such as Rodriguez and Rodrik (1999) questioning the earlier studies that showed that trade liberalization had led to faster growth were also given less play.

28. Sometimes the reason for this was closely related to the reason for advocating the reforms. Economists typically advocate greater transparency, partly because bringing certain costs—for instance, those of subsidies—into the open will erode support for them. But the beneficiaries of these subsidies also recognized that increased transparency would threaten the sustainability of the subsidies.

29. See Stiglitz (1999d) and Roland (2000), a textbook on transition economics that summarizes the literature on this point and advances a critique of the "cavalry approach" to transition.

30. Indeed, the irony is that the ascendancy of the neoliberal doctrines based on the neoclassical model occurred at a time that economic theory was stressing the limitations of that model and developing an alternative paradigm based on imperfect information and incomplete markets (and associated other market failures, including imperfections of information).

31. See Wolfensohn (1998) and www.worldbank.org/cdf.

References

- Abreu, Dilip. 1988. "On the Theory of Infinitely Repeated Games with Discounting." *Econometrica* 56 (2): 383–96.
- Aghion, Philippe, and Patrick Bolton. 1997. "A Theory of Trickle-Down Growth and Development." *Review of Economic Studies* 64 (2): 151–72.

- Akerlof, George A. 1985. "Discriminatory, Status-Based Wages among Tradition-Oriented, Stochastically Trading Coconut Producers." *Journal of Political Economy* 93 (2): 265–76.
- Arnott, R., and Joseph E. Stiglitz. 1991. "Moral Hazard and Non-Market Institutions: Dysfunctional Crowding Out or Peer Monitoring?" *American Economic Review* 81 (1): 179–90.
- Banerjee, Abhijit, D. Mookherjee, K. Munshi, and D. Ray. 2001. "Inequality, Control Rights and Rent Seeking: Sugar Cooperatives in Maharashtra." *Journal of Political Economy* 109 (1): 138–90.
- Bardhan, P. K. 1989. *The Economic Theory of Agrarian Institutions*. New York: Oxford University Press.
- Basu, Kaushik, Garance Genicot, and Joseph Stiglitz. 1999. "Household Labor Supply, Unemployment, and Minimum Wage Legislation." Policy Research Working Paper 2049. World Bank, Washington, D.C.
- Becker, Gary S. 1957. *The Economics of Discrimination*, Chicago: University of Chicago Press.
- Bigsten, Arne, and Karl Ove Moene. 1997. "Growth and Rent Dissipation: The Case of Kenya." *Journal of African Economies* 5 (2): 177–98.
- Birdsall, Nancy, Carol Graham, and Richard H. Sabot, eds. 1998. *Beyond Tradeoffs: Market Reforms and Equitable Growth in Latin America*. Washington, D.C.: Inter-American Development Bank and Brookings Institution Press.
- Black, Bernard, Reinier Kraakman, and Anna Tarassova. 2000. "Russian Privatization and Corporate Governance: What Went Wrong?" *Stanford Law Review* 52 (6): 1731–1808.
- Bourguignon, Francois, and T. Verdier. 2000. "Is Financial Openness Bad for Education? A Political Economy Perspective on Development." *European Economic Review* 44 (4–6): 891–903.
- Cheung, Steven. 1969. *The Theory of Share Tenancy*. Chicago: University of Chicago Press.
- Coase, R. H. 1992. "The Institutional Structure of Production." *American Economic Review* 82 (4): 713–19.
- Darwin, Charles. [1859] 1993. *The Origin of Species by Means of Natural Selection*. Modern Library edition. New York: Random House.
- Dasgupta, Partha, and Ismail Serageldin. 2000. *Social Capital: A Multifaceted Perspective*. Washington, D.C.: World Bank.
- Dewey, John. 1939. *Freedom and Culture*. New York: Capricorn.
- Furman, Jason, and Joseph E. Stiglitz. 1998. "Economic Crises: Evidence and Insights from East Asia." *Brookings Papers on Economic Activity* 2: 1–114.
- Greenwald, Bruce, and Joseph E. Stiglitz. 1986. "Externalities in Economies with Imperfect Information and Incomplete Markets." *Quarterly Journal of Economics* 101 (2): 229–64.
- Haas, E. B. 1990. *When Knowledge Is Power: Three Models of Change in International Organizations*. Berkeley: University of California Press.
- Hirschman, Albert O. 1981. *Essays in Trespassing: Economics to Politics and Beyond*. Cambridge: Cambridge University Press.
- Hoff, Karla. 1994. "The Second Theorem of the Second Best." *Journal of Public Economics* 54 (2): 223–42.
- Hoff, Karla, and Joseph E. Stiglitz. 2001. "Modern Economic Theory and Development." In Gerald Meier and J.E. Stiglitz, eds., *The Future of Development Economics in Perspective*. Oxford: Oxford University Press.
- Hoff, Karla, Avishay Braverman, and Joseph E. Stiglitz, eds. 1993. *The Economics of Rural Organization: Theory, Practice, and Policy*. New York: Oxford University Press.
- Hussein, Athar, Nicholas Stern, and Joseph E. Stiglitz. 2000. "Chinese Reforms from a Comparative Perspective." In Peter J. Hammond and Gareth D. Myles, eds., *Incentives*,

- Organization, and Public Economics: Papers in Honour of Sir James Mirrlees*. New York: Oxford University Press.
- Kaldor, Nicholas. 1956. "Alternative Theories of Distribution." *Review of Economic Studies* 23 (2): 83–100.
- Krueger, Anne O. 1974. "The Political Economy of the Rent-Seeking Society." *American Economic Review* 64 (3): 291–303.
- Kuznets, Simon. 1955. "Economic Growth and Income Inequality." *American Economic Review* 45 (1): 1–28.
- Lewis, W. A. 1954. "Economic Development with Unlimited Supplies of Labor." *Manchester School of Economic Studies* 22 (May): 139–91.
- Lucas, Robert E. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (1): 3–42.
- Marshall, Alfred. 1897. "The Old Generation of Economists and the New." *Quarterly Journal of Economics* 11 (2): 115–35.
- Morduch, Jonathan. 1999. "The Microfinance Promise." *Journal of Economic Literature* 37 (4): 1569–1614.
- Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny. 1989. "Industrialization and the Big Push." *Journal of Political Economy* 97 (5): 1003–26.
- Myrdal, Gunnar. 1957. *Economic Theory and Underdeveloped Regions*. New York: Harper Torchbooks.
- . 1968. *Asian Drama: An Inquiry into the Poverty of Nations*. 3 vols. New York: Pantheon.
- North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.
- Orszag, Peter R., and Joseph E. Stiglitz. 2000. "Rethinking Pension Reform: Ten Myths About Social Security Systems." In Robert Holzmann and Joseph E. Stiglitz, eds., *New Ideas about Old Age Security*. Washington, D.C.: World Bank.
- Osborne, David, and Ted Gaebler. 1992. *Reinventing Government: How the Entrepreneurial Spirit Is Transforming the Public Sector*. Reading, Mass.: Addison-Wesley.
- Peyton, Young. 1996. "The Economics of Convention." *Journal of Economic Perspectives* 10 (2): 105–22.
- Piketty, Thomas. 1997. "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing." *Review of Economic Studies* 64: 173–89.
- Platteau, Jean-Philippe. 1994. "Behind the Market Stage Where Real Societies Exist—Part II: The Role of Moral Norms." *Journal of Development Studies* 30 (4): 753–817.
- Polanyi, Karl. 1944. *The Great Transformation: The Political and Economic Origins of Our Time*. New York: Rinehart and Company.
- Putnam, Robert. 1993. *Making Democracy Work*. Princeton, N.J.: Princeton University Press.
- Rodriguez, Francisco, and Dani Rodrik. 1999. "Trade Policy and Economic Growth: A Skeptic's Guide to Cross-National Evidence." NBER Working Paper W7081. National Bureau of Economic Research, Cambridge, Mass.
- Rodrik, Dani. 1998. "Who Needs Capital-Account Convertibility?" *Essays in International Finance* 207: 55–65.
- Roland, Gérard. 2000. *Transition and Economics: Politics, Markets and Firms*. Cambridge, Mass.: MIT Press.
- Rosenstein-Rodan, Paul. 1943. "Problems of Industrialization of Eastern and South-Eastern Europe." *Economic Journal* 53 (June–September): 202–11.
- Rosenzweig, Mark R. 1988. "Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries." *Economic Journal* 98 (393): 1148–70.

- Schön, Donald A. 1983. *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.
- Shleifer, Andrei, and Robert W. Vishny. 1998. *The Grabbing Hand: Government Pathologies and Their Cures*. Cambridge, Mass.: Harvard University Press.
- Simon, Herbert A. 1991. "Organizations and Markets." *Journal of Economic Perspectives* 5 (2): 25–44.
- Stiglitz, Joseph E. 1974a. "Incentives and Risk Sharing in Sharecropping." *Review of Economic Studies* 41 (2): 219–55.
- . 1974b. "Theories of Discrimination and Economic Policy." In George M. von Furstenberg, Bennett Harrison, and Ann R. Horowitz, eds., *Patterns of Racial Discrimination*. Lexington, Mass.: Lexington Books.
- . 1987a. "On the Microeconomics of Technical Progress." In Jorge M. Katz, ed., *Technology Generation in Latin American Manufacturing Industries*. Basingstoke: Macmillan.
- . 1987b. "Technological Change, Sunk Costs, and Competition." *Brookings Papers on Economic Activity* 3 (0): 883–937.
- . 1988. "Economic Organization, Information, and Development." In Hollis Chenery and T.N. Srinivasan, eds., *Handbook of Development Economics*. New York: Elsevier Science Publishers.
- . 1989a. "The Economic Role of the State." In Arnold Heertje, ed., *The Economic Role of the State*. Oxford: Basil Blackwell in association with Bank Insinger de Beaufort NV.
- . 1989b. "Rational Peasants, Efficient Institutions and the Theory of Rural Organization." In P.K. Bardhan, ed., *The Economic Theory of Agrarian Institutions*. Oxford: Clarendon Press.
- . 1990. "Peer Monitoring and Credit Markets." *World Bank Economic Review* 4 (3): 351–66.
- . 1991a. "Development Strategies: The Roles of the State and the Private Sector." In *Proceedings of the World Bank Annual Conference on Development Economics 1990*. Washington, D.C.: World Bank.
- . 1991b. "The Economic Role of the State: Efficiency and Effectiveness in the Public Domain." In T. P. Hardiman and M. Mulreany, eds., *Efficiency and Effectiveness in the Public Domain*. Dublin: Institute of Public Administration.
- . 1992. "Contract Theory and Macroeconomic Fluctuations." In L. Werin and H. Wijkander, eds., *Contract Economics*. Oxford: Basil Blackwell.
- . 1994. *Whither Socialism?* Cambridge, Mass.: MIT Press.
- . 1995a. "Social Absorption Capability and Innovation." In Bon Ho Koo and D. H. Perkins, eds., *Social Capability and Long-Term Economic Growth*. New York: St. Martin's Press.
- . 1995b. "The Theory of International Public Goods and the Architecture of International Organizations." Background Paper 7. Third Meeting, High-Level Group on Development Strategy and Management of the Market Economy, United Nations University, World Institute for Development Economics Research, Helsinki, 8–10 July.
- . 1996a. "The Role of the Government in the Economy of Less Developed Countries." In István P. Székely and Richard Sabot, eds., *Development Strategy and Management of the Market Economy*. New York: United Nations.
- . 1996b. "Some Lessons from the East Asian Miracle." *World Bank Research Observer* 11 (2): 151–77.
- . 1998a. *More Instruments and Broader Goals: Moving Toward the Post-Washington Consensus*. World Institute for Development Economics Research (WIDER) Annual Lectures 2. Helsinki: WIDER.

- . 1998b. “The Role of Government in Economic Development.” In Michael Bruno and Boris Pleskovic, eds., *Annual World Bank Conference on Development Economics 1996*. Washington, D.C.: World Bank.
- . 1998c. “Towards a New Paradigm for Development: Strategies, Policies, and Processes.” Paper presented as the Prebisch Lecture at the United Nations Conference on Trade and Development, Geneva.
- . 1998d. “The Private Uses of Public Interests: Incentives and Institutions.” Distinguished Lecture on Economics in Government. *Journal of Economic Perspectives* 12 (2): 3–22.
- . 1999a. “Knowledge as a Global Public Good.” In I. Kaul, I. Grunberg, and M. A. Stern, eds., *Global Public Goods—International Co-operation in the 21st Century*. New York: Oxford University Press for the United Nations Development Programme.
- . 1999b. “Knowledge for Development: Economic Science, Economic Policy, and Economic Advice.” In Boris Pleskovic and Joseph E. Stiglitz, eds., *Annual World Bank Conference on Development Economics 1998*. Washington, D.C.: World Bank.
- . 1999c. “Quis Custodiet Ipsos Custodes? (Who Is to Guard the Guards Themselves?) Corporate Governance Failures in the Transition.” *Challenge* 42 (6): 26–67.
- . 1999d. “Whither Reform? Ten Years of the Transition.” *Transition Economics* 3 (12).
- . 2000a. “Formal and Informal Institutions.” In P. Dasgupta and I. Serageldin, eds., *Social Capital: A Multifaceted Perspective*. Washington, D.C.: World Bank.
- . 2000b. “Scan Globally, Reinvent Locally: Knowledge Infrastructure and the Localization of Knowledge.” In Diane Stone, ed., *Banking on Knowledge*. New York: Routledge Press.
- Stiglitz, Joseph E., and Lyn Squire. 1998. “International Development: Is It Possible?” *Foreign Policy* 110 (spring): 138–51.
- Wolfensohn, James D. 1996. “People and Development.” *Annual Meetings Address*. Washington, D.C.: World Bank.
- . 1998. “The Other Crisis.” *Address to the Board of Governors*. Washington, D.C.: World Bank.
- World Bank. 1993. *The East Asian Miracle: Economic Growth and Public Policy*. New York: Oxford University Press.
- . 1998. *Assessing Aid: What Works, What Doesn't, and Why*. New York: Oxford University Press.
- . 1999. *World Development Report 1998/99: Knowledge for Development*. New York: Oxford University Press.
- . 2000. *World Development Report 2000/2001: Attacking Poverty*. New York: Oxford University Press.

KEYNOTE ADDRESS

A New Global Consensus on Helping the Poorest of the Poor

Jeffrey Sachs

There is an urgent need for a new consensus on economic development. The World Bank and the International Monetary Fund (IMF) should take notice when many of their natural supporters have become their ardent opponents—though much of the criticism is misplaced. The Bretton Woods institutions are bearing the brunt of the fact that rich countries, especially the United States, have largely turned their backs on the world's poorest people. But the Bretton Woods institutions have been willing accomplices in the dismantling of an effective agenda for global poverty alleviation. Because these institutions are owned and operated by their shareholders—with a clear majority held by the United States and Western Europe—the IMF and World Bank have defended the ever-shrinking and unrealistic development agenda. To do otherwise would insult the leading shareholders, who pay the bills and choose the management.

Rich Countries, Poor Aid—and Disastrous Results

The U.S. position is clear enough: “We already paid during the cold war, so leave us alone and let us enjoy our wealth and new economy.” Is this an unfair characterization? No. Beneath the high-minded rhetoric of a high-minded administration lies a grim reality. In 1998 U.S. foreign aid totaled \$8.8 billion, or 0.11 percent of gross national product (GNP). And of this derisory sum, only a sixth went to the world's least developed countries. One-sixth of eleven-hundredths of one percent of GNP amounted to a grand total of about \$5 per American for the least developed countries. This is \$5 a year in a country where the average income is more than \$30,000, and where investors have enjoyed more than \$7 trillion in capital gains since the beginning of 1996.

Consider it from the recipient side. In 1998 there were about 600 million people in the least developed countries, so U.S. aid amounted to \$2.20 a person. Looking

Jeffrey Sachs is director of the Center for International Development at Harvard University and Galen L. Stone Professor of International Trade at Harvard University.

Annual World Bank Conference on Development Economics 1999

©2000 The International Bank for Reconstruction and Development / THE WORLD BANK

not just at the least developed countries but at all low-income countries (defined as those with GNP per capita below \$765 in 1995), 3.36 billion people received \$3.7 billion in U.S. aid, or \$1.13 a person.

I said that the U.S. administration is high-minded. I have no doubt that President Clinton and Secretary of Treasury Lawrence Summers have great concern for the world's poor. But they apparently feel that they cannot act much on that concern. The \$5 per American is not the result of Congress cutting a large portion of the president's aid request—it basically is the president's aid request. Yet the administration came up with \$8 billion to fight the 79-day war in Kosovo, and \$1.6 billion for 30 Blackhawk helicopters to fight the drug war in Colombia, and it routinely comes up with billions for the Middle East. But how much has the administration requested from Congress for Ethiopian famine relief or for long-term improvements in Ethiopian agriculture? And how much has the administration requested for Nigeria, the most populous and economically important country in tropical Africa, now with an unprecedented but uncertain opportunity to make democracy work?

Senior U.S. officials have told me repeatedly that the administration feels unable to ask Congress for even the \$150 million that it would cost to forgive the more than \$1 billion that Nigeria owes to the United States. Yet Nigeria is so bankrupt that its arrears on foreign debt are nearly 40 percent of GDP, and it spends five times more servicing debt than it spends on public health.

The Clinton administration seems to feel that it cannot do more. I strongly disagree. As I will explain, the administration could mobilize much greater sums as part of a revised global strategy to fight global poverty. But it has never presented such a strategy to the American people to put that proposition to the test.

What are the results of the world's minimalist approach to helping the poor? International efforts to alleviate poverty are profoundly underfunded and consequently half-baked. A case in point is the program to relieve the debt of heavily indebted poor countries—the so-called HIPC initiative. The program has been so badly mangled by the international community that millions of people have protested the debt relief policies that the IMF and World Bank consider their finest hour.

Indeed, everywhere one turns in global poverty reduction efforts, high-minded rhetoric provides a tattered veneer over deficient funding. The AIDS epidemic has flared in the world's poorest countries, especially in Sub-Saharan Africa, for more than 10 years. Yet rich countries and the Bretton Woods institutions have put almost no money toward battling it. World Bank President James Wolfensohn has done as much as anyone in recent years to bring AIDS to the attention of the global community. But how much has the International Development Association (IDA) lent for AIDS over the past 20 years, from the time the epidemic was getting under way to the more than 33 million people infected today, and more than 16 million deaths? According to a 1999 World Bank report, IDA devoted \$340 million to AIDS in 1986–98, or about \$26 million a year. That comes out to around 4 cents an African each year. In short, the Bank, and the donor community more broadly, have stayed on the sidelines in the face of the worst epidemic in modern history.

The situation is no better for malaria, another killer that claims at least 1 million lives a year and causes up to half a billion clinical cases a year. The World Bank recently confirmed earlier findings by Harvard's Center for International Development that malaria not only takes lives but also cripples economic growth—lowering growth by at least 1 percentage point a year in the hardest-hit countries. So where are the Bank's malaria projects in Africa? There are almost no standalone projects. Malaria control has collapsed in Africa, and it will require at least \$1 billion a year to get malaria back under control. I will stress this point at next week's African Summit on Malaria, convened by Nigerian President Olesgun Obasanjo in conjunction with the World Health Organization. Will the Bank and other donors hear the message?

Similar tragedies are played out every day in IMF lending programs in the poorest countries. The IMF starts with the truth that budget deficits should stay small to preserve macroeconomic stability. Then it demands budget austerity of impoverished countries to the point where those countries cannot even keep their people alive, so depleted are the budgets for public health, food transfers to the poor, and the like. In addition, the IMF has repeatedly insisted on debt servicing that exceeds the combined spending of health and education ministries.

Yet when the world complains about the disasters of IMF loan conditions, the IMF's response is that the protestors are obviously macroeconomic illiterates. I am not a macroeconomic illiterate, and I can tell you that the budget conditions in the world's poorest countries are unconscionable. These countries need a lot more help. Yes they should balance their budgets, but in a context of much greater aid with cancellation of their debts. The IMF should trumpet this truth, not hide it.

The Escape from Poverty—Relying on More than Economic Policies

At least since the early 1980s, both Democratic and Republican administrations in the United States have offered an ideological fig leaf for the tragic underfunding of poverty reduction programs. This has since become the mantra of the IMF and World Bank, and it goes something like this:

Poverty reduction is mainly the result of economic growth, which in turn is mainly the result of good economic policies. Nothing that blocks economic development in Burkina Faso or Ethiopia or Nepal cannot be fixed through effective economic policies centered on macroeconomic stability, open trade and finance, government support for social programs, and privatization. If poverty is not falling, it is the result of poor governance, in the sense that one or more of those reform items remains unfulfilled. The IMF and World Bank, together with partial debt relief, can play a modest role in filling the financing needs of countries while they make needed policy adjustments, but Washington cannot substitute for good governance or overcome corruption.

This simple-minded reasoning is based vastly more on convenience than on evidence or analysis. It is a fancy way of telling poor countries not to come to us with

their problems, and certainly not to ask for more financial help. Economic reforms are certainly important, and I have spent the past 15 years helping dozens of countries implement such reforms. But these reforms are only part of the story, and for many of the poorest countries they are not the most important explanation for their continuing desperation and impoverishment.

The escape from poverty rests on four pillars, not just economic reform. A second pillar is having a population that is healthy and educated enough to participate in the global economy. Many of the world's poorest places are too sick and too lacking in education to succeed. Life expectancy is often 50 years or less, and is plummeting in much of Africa because of AIDS. About 40 percent of the children in heavily indebted poor countries are malnourished. Adverse climatic and agronomic conditions often impose barriers that earlier reformers such as the Republic of Korea and Taiwan (China), or the United States for that matter, did not face—such as holoendemic falciparum malaria, degraded tropical soils, or extremely scarce clean water in deserts, steppes, and tropical savannahs.

The poorest countries lack the resources to overcome these hurdles on their own. At an income of \$300 per capita, even a budget outlay of 5 percent of GDP for public health—much more than in almost any developing country—is just \$15 per person per year, a sum clearly insufficient to meet basic health needs.

The third pillar of development is technology. The fuel of U.S. prosperity, as President Clinton would be the first to note, is technological growth. And despite the free-market rhetoric of the United States, technological change is the product of a complex system of private, public, and academic institutions, and the financing comes from markets, government, and foundations. It is no coincidence that each year the “free market” United States spends \$85 billion in public funds on basic science and applied research and development. But what of technological development in the poorest countries, to meet their specific needs—for a malaria or HIV/AIDS vaccine, or for enhanced crops that can withstand salinization of irrigated land or heat and drought stress, or for new forms of energy that can slow tropical deforestation? Add up all the World Bank grants and loans for science and technology for all poor countries last year, and I bet that it is less than a fifth of the research and development budget of a single major U.S. pharmaceutical company. I would be grateful if someone at the Bank would check this bet.

The fourth pillar of poverty reduction is structural adjustment, especially export diversification. Here too rich country convenience, sheltered by ideology, intrudes on the needs of the poorest countries. Over the past 20 years structural adjustment has become a detested phrase among antipoverty activists. But that is because World Bank structural adjustment programs have often been the opposite of structural adjustment. Sub-Saharan African countries are as dependent on a narrow range of primary commodities today as they were 20 years ago, but now real world prices for those commodities are even lower. In fact, the Bank has usually acted as if there is no need to foster manufactured exports from Africa, content to encourage greater reliance on primary commodities. True structural adjustment requires a strategy to foster new kinds of industry, and it requires open markets in the United States and

Europe for manufactured exports from the poorest countries, especially in textiles and garments.

One of my great frustrations is that the World Bank has been one of the leading opponents of export zones, tax holidays, and other basic industrial policies that have been key to success elsewhere, such as in the East Asian miracle countries or in Boston for that matter. It is no accident that U.S. ecommerce has been nurtured by a tax holiday. The United States even gives extensive tax breaks to exporters. When it comes to industrial policies, trade policies, and technology policies, the rich country and Bretton Woods position is "Do as I say, not as I do."

These four pillars may seem like truisms: that economic reform must be combined with enough resources to meet basic human needs; that priority should be given to developing new health, agriculture, and energy technologies that are ecologically specific and where rich country technologies will not suffice; and that countries should pursue industrial policies geared toward diversifying away from dependence on bananas, coffee, tea, minerals, and other primary commodities, and that such policies should be supported by market access in rich countries.

So if it is obvious, why does this more complete agenda not get heard, much less implemented? A small part of the answer has to do with ideology. Extreme free marketeers might object to the idea that market reforms are insufficient or that governments should have technology and industrial policies of any kind. But most of the participants in this discussion in the U.S. government, the IMF, and the World Bank are not extremists. There are two simpler reasons for the unfulfilled agenda.

First, for a large and growing part of the world, the four pillars are coming into place on their own. I do not think that we really have to worry about Poland's capacity to grow if the European Union carries through on its pledges to expand. The same is basically true of the Baltics, Croatia, the Czech Republic, Hungary, Slovakia, and Slovenia. Mexico will grow quickly under the North American Free Trade Agreement if political liberalization stays on course. So will Egypt, Morocco, and Tunisia under the Mediterranean Agreements with the European Union, though Egypt's problems are complicated by demographic and environmental stress.

The main problems I am emphasizing are concentrated in the poorest parts of the world—Sub-Saharan Africa, much of the Andean region, the Gangetic valley of India, Central Asia, parts of western China (if growth in coastal China is not strong enough)—where geographic isolation, difficult climate, disease, and mass illiteracy are too overwhelming to be solved by adjusting economic policies. Do not think that because globalization is working powerfully for some regions that it is working powerfully for all regions.

But even this point would be better understood were it not for the second, more important, reason for the unfulfilled poverty reduction agenda. Such an agenda would cost money, a lot more than is now being offered by rich countries. It would also require facing up to U.S. and European protectionist lobbies, which fight imports of garments and other assembled goods from Africa and Asia, and to the imbalances of a global trading system increasingly subdivided by regional trade agreements that discriminate against poor countries on the geographic margins of

the world. And of course, it would require us to think harder, to move beyond the easy platitudes claiming good governance in the poorest countries as the solution to all problems.

Moving beyond Inaction

I have watched with fascination in recent months as the insufficiencies of the global poverty agenda have been debated. The United States, Europe, and the Bretton Woods institutions seem trapped in inaction even as the inadequacies of the current situation bring thousands of demonstrators to the streets, albeit often in a confused and unsatisfying manner. But the bottom line of many demonstrators is completely right: the current situation condemns hundreds of millions of people to unnecessary suffering and millions to premature death, and the Bretton Woods institutions are parties to the disaster. Ultimate responsibility falls squarely on the leading shareholders of the institutions, especially the United States, but the IMF and World Bank have been willing accomplices by lending their names to grossly underfunded and insufficient strategies.

Let me put one issue to rest. I do not agree with my brave and brilliant friend Joseph Stiglitz, who recently characterized IMF staff as third rank. I know the staff to be first rate in dedication and, I might say, education, since many have been my most prized students. I hope they feel the same way about their former teacher. In addition, Bank President James Wolfensohn and IMF Acting Managing Director Stanley Fischer are men of world-class intelligence, energy, and integrity, and so is the incoming head of the IMF, Horst Kohler. But all these people are operating in a system that is thoroughly unsatisfactory for the world's poorest countries, those caught in the vice of disease, geographic isolation, illiteracy, and impoverishment. With all due respect, I believe that the managers and staffs of the IMF and World Bank have too often defended that system without showing the world how tragically unnecessary the extreme suffering really is.

The system persists in an interlocking series of excuses. U.S. Treasury Secretary Lawrence Summers has said that he would like the United States to do more but that Congress would block such efforts. That may be true, but as noted the Clinton administration has never offered an ambitious new approach to international aid, based on increased funding and fundamental reforms of aid delivery, to see whether Americans would support it. The IMF and World Bank have occasionally argued for additional debt relief and assistance, but claimed that there are not sufficient funds available. But the IMF and Bank could take the lead on debt cancellation by writing off their Enhanced Structural Adjustment Facility and IDA credits without damaging IMF quotas or Bank capital. The U.S. Congress has understandably opposed additional funds for international institutions because they have failed to deliver the benefits they have repeatedly promised, but Congress has never said that more funding would ensue in the event of reform.

The street protestors see the mess and condemn the system. In response, the U.S. government, IMF, and Bank dig in their heels against what they consider a benighted

mob, and make unjustified claims about all the good they are doing. Rather than admitting that their hands are tied by the lack of resources from rich countries, the IMF and Bank defend their shareholders. The bad will and misunderstanding simply cascade.

The Meltzer Commission, on which I served, assessed the role of international institutions and gave more than a hint of how to break the deadlock. Broad bipartisan consensus within the commission indicates that it would be possible to mobilize much greater U.S. assistance for the poorest countries as part of a revived and revised strategy for global poverty alleviation. A bipartisan approach could work in Congress and the country at large.

In my opinion, and that of the commission, this strategy should have five elements. First, admit the obvious: the world's poorest countries need much more help than is being offered. To address the health crisis alone would require several billion dollars more each year. I am delighted that the World Bank and IMF are working closely with the World Health Organization as part of its Commission on Macroeconomics and Health, which I chair, to realistically assess global needs in public health. We will be reporting on those needs next year.

Second, recognize that this help should come in new ways. Technological development—such as a malaria vaccine—will require major grants to science-based institutions, as well as new partnerships between business and academia spurred by innovative institutional arrangements. The Clinton administration has endorsed the idea of a tax credit or guaranteed purchase fund to spur research and development for new vaccines. The Bank should commit IDA funds to this proposal as well. Traditional Bank loans to countries are almost certainly the wrong way to encourage needed technologies, but country programs can improve the public health systems that will be needed to deliver those technologies.

Third, get institutions back to their relevant roles. The IMF finds itself deep in African development for artificial reasons or reasons that are now passé. The Bank often takes the lead on an underfunded global health agenda because the World Health Organization, like other United Nations agencies, is even more squeezed for cash. The IMF should simply get out of poverty lending, a view endorsed 11–0 by the Meltzer Commission, with the IMF's concessional money being transferred through other agencies—especially the World Health Organization, United Nations Children's Fund, and World Bank. Stanley Fischer suggests that every poor IMF member has the right to IMF concessional lending. But the issue is not rights, but the effectiveness of global assistance. Both the IMF as an institution and the world's poverty relief efforts have been damaged by the IMF's improper role in development lending.

Fourth, release heavily indebted poor countries from their misery once and for all by fully canceling their debts, not going halfway as in the current initiative. This was another unanimous recommendation of the Meltzer Commission. Current debt reduction targets are based on a phony debt sustainability analysis that could not pass muster in a first-year economics class. Indeed, the phrase "debt sustainability analysis" is Orwellian in its scale of distortion. IMF and World Bank procedures for measuring debt sustainability have nothing to do with a country's ability to pay, and

everything to do with the arbitrary limits on debt relief laid down by the Group of Seven (G-7) rich countries. IMF and World Bank documents should be relabeled as “debt relief allowed by the G-7” rather than “debt sustainability analysis.” At least the world would complain less about the roles of the IMF and Bank in this sham, and turn the spotlight on creditor countries instead.

Fifth, as the counterpart to greatly increased funding for and focus on the poorest countries, there should be a recognition inside the Bank that countries like Argentina, Brazil, Chile, Republic of Korea, and Mexico are not the proper focus of Bank lending. These countries absorb a large portion of Bank time and attention—not to mention loans—and they distract from the much harder work of solving the problems of the world’s poorest countries, as well as global problems such as human-induced climate change.

A strategy like this—focusing on the poorest countries, taking new approaches to technological development, getting the IMF back to its core business, canceling unpayable debts, and getting the Bank to scale back its activities in richer countries while scaling up its support for the poorest countries and for global public goods—would win broad bipartisan approval. Even conservative members of Congress would sign an effective assistance strategy that delivers real benefits to the world’s poorest people, especially for programs (such as vaccine development) that create knowledge to alleviate poverty. In my experience, opposition to foreign aid has intensified in recent years because it is viewed as a failure, not out of cold-heartedness to the plight of the world’s neediest people.

With an expanded aid budget, the global community could do wonderful things for poor countries. Rather than pursue limited flows of Bretton Woods country lending with harsh conditions attached, the world could support breakthroughs in health and agricultural research that would make a difference in the long run. Rather than have extremely limited health programs, the World Health Organization could once again take the lead in identifying and targeting the interventions needed to cure malaria, tuberculosis, diarrheal disease, and AIDS, as it has with smallpox and now nearly with polio. Rather than pursuing specific disease research programs whenever the World Health Organization and World Bank get the funds for them, a network of health research institutions could be created around the developing world to pursue this research in earnest. This Consultative Group for International Health Research could complement the Consultative Group for International Agriculture Research. Combining information technology with breakthroughs in biotechnology and other areas creates immense possibilities for progress in health, agriculture, and environmental management.

For much of the 20th century the Rockefeller Foundation showed the world what grants focused on knowledge could do. Rockefeller funds supported the eradication of hookworm in the U.S. South, the discovery of the Yellow Fever vaccine, the accelerated development of penicillin, the control of malaria in Brazil, the establishment of leading public health schools and medical facilities all over the world, the establishment of research centers that drove the green revolution in Asia and that became the Consultative Group for International Agricultural Research, and the establish-

ment and funding of great research centers such as the University of Chicago, Brookings Institution, Rockefeller University, and National Bureau of Economic Research.

None of these earth-shaking accomplishments was the result of a high-conditionality country loan. All required large-scale grants ready to back the pursuit of knowledge. Indeed, the donor wanted to build strong and independent institutions, so the Rockefeller Foundation consciously and explicitly eschewed conditionality.

In our own time, the Bill and Melinda Gates Foundation has taken a similarly bold tack, with major new support for public health initiatives and institution building—most notably for delivering vaccines to poor countries through the creation and financing of a new Global Alliance for Vaccines and Immunization.

My colleagues on the Meltzer Commission and I believe that the global community should consider these remarkable examples as guides in reforming the Bretton Woods process. The kind of help that poor countries need is so different from the World Bank's traditional lending that the Bank should highlight its retreat from banking and its refocus on knowledge creation by changing its name from the World Bank to the World Development Agency. The world has thousands of banks, but it desperately needs an institution charged with creating and mobilizing knowledge for development.

I have fought hard to reform the IMF and World Bank not because I am a foe of these institutions, but because I am a strong supporter. I believe in the quaint concept of a global community, and I believe in shared global governance. But I do not believe in global governance by rich countries, or international voting weighted by money as in the IMF and Bank, or permanent government by entrenched bureaucracies unencumbered by external review as has been true of the IMF, or governance through conditionality set by rich countries and imposed on the desperately poor.

It is time for the World Bank and the IMF to assert their intellectual leadership and independence and to show the world the greatly increased and urgent efforts that must be made on behalf of the world's poorest people. I know that you can do it, and I would be proud to work with you on that valuable task.

KEYNOTE ADDRESS

Ten Years After *The Road to a Free Economy*: The Author's Self-Evaluation

János Kornai

Ten years have passed since the publication of my book *The Road to a Free Economy: Shifting from a Socialist System—The Example of Hungary*. It was the first book to offer comprehensive proposals for the postsocialist transition. This article assesses the book as I see it 10 years later.¹ Is this not an extremely self-centered undertaking? An advertisement for an old book that no one is buying these days? No. There are good ethical and intellectual reasons for assessing the book, as will be made clear.

Customary measures of success in the academic world try to measure a work's impact on its author's colleagues. Here I could be satisfied. Several hundred references have been made to the book—including by scholars who disagreed with what I said. Authors are gratified if their work proves controversial.

But here citations are not a sufficient measure of success. Because the book offered policy recommendations, a far more serious question has to be asked: what impact did it have on the outside world? I am not like a meteorologist, who makes a forecast but then sees the weather develop of its own accord. When I wrote *The Road to a Free Economy*, I expected it to have at least a modest impact on public opinion and political decisions, and ultimately to influence the course of events.

History is not shaped by blind forces. It is influenced by conscious people who bear responsibility for their actions. The main historical responsibility falls on politicians, but in the second rank stand advisers from the academic world. They too are accountable for what they say.²

Heated debates broke out in the early 1990s on what strategy should be adopted for the transition. (For a summary, see Roland 2000, especially chapters 4 and 10.) I will return to those debates, though not in a combative form. I will contrast my

János Kornai is professor of economics at Harvard University and Collegium Budapest. An earlier version of this article was delivered as the keynote address to the Nobel Symposium held in Stockholm in September 1999. The author is grateful to the symposium participants and to Zsuzsa Dániel, Stanislaw Gomulka, Karel Kouba, and Kazimir Poznanski for stimulating comments and suggestions, to Mária Barát, Ágnes Benedict, Andrea Despot, Cecília Hornok, and Julianna Parti for efficient research assistance, and to Brian McLean for excellent translation.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

views with those of others, but without pointing fingers. I hope that this approach will prevent today's debates from becoming personal and direct attention to the problems themselves.

The focus here is on self-evaluation. I will avoid self-justification and self-congratulation, and I will be self-critical. But modesty will not keep me from endorsing my earlier views if I believe they are still legitimate.

How can it be established whether the book's messages were right or wrong? It is not enough simply to compare it with the facts. A case where the course of events coincided with my advice could be unfortunate if my recommendations were wrong. And it could be fortunate that events did not coincide with my recommendations if they were mistaken.

Whatever approach is used to judge the recommendations, the real task is to assess the events themselves—the actual course of history. That cannot be done without making value judgments. I will refrain from stating in advance the values through which I view the events, but will reveal them step by step. Ultimately, my conscience is the judge.

My book was originally written for the Hungarian public. The Hungarian edition appeared in 1989, before the country's first free parliamentary elections, and eventually appeared in 16 other languages with minor alterations. The foreword to the foreign editions warned that the recommendations could not be applied mechanically to other countries. Although I considered many aspects of the recommendations to have universal validity, they needed to be adapted to conditions in each country. So it seems expedient to focus here on Hungary's experience, augmenting it with references to developments in the Czech Republic, Poland, and Russia.

A complete account would have to cover the 15–20 issues discussed in the book. In hindsight, I see that I was right on many but wrong on quite a few others. I hope that someday I will have a chance to make a more detailed assessment, but here I will confine myself to just two issues.

The first issue is ownership reform and private sector development, where I believe that my recommendations were fundamentally correct. The second is macroeconomic stabilization. Here my report card is mixed: I was partly right and partly wrong.

Ownership Reform and Private Sector Development

The Road to a Free Economy took issue with the basic concept of market socialism. It rejected the idea that dominant state ownership should be retained but linked to market coordination. This position irritated advocates of market socialism—particularly many reform economists in Eastern Europe and many old-style social democrats in the West.

The book reflected its author's credo, supporting an economic system in which private ownership would dominate. In this respect the book did not differ from many proposals originating in the West. But this broad agreement left open important questions. What is the best road to such a system? Once the transition is over,

what will be the economy's ownership structure? Of the many variants of capitalism based on private ownership, which one should be sought?

Many ideas arose for private sector development. Here I briefly set out two pure strategies. Most of the detailed, practical proposals of the early 1990s came close to one or the other, and conflict between them lay at the center of the debates.

The Organic Development Strategy

The first strategy, which I call the strategy of organic development, has five main features:

- The most important task is to create favorable conditions for bottom-up development of the private sector. Mass entry of new firms is the main impetus for private sector growth. This process must be assisted by breaking down barriers to free entry, by guaranteeing the security of private ownership (with institutions created to enforce private contracts), and by applying “affirmative action”—with caution—to promote private sector development (as in tax and credit policy).
- Most state enterprises must be privatized. Sales are the basic technique for doing so. State assets have to be sold mainly to outsiders, giving preference to those who are willing to pay a fair price and commit to investing in the company. If the buyer is an insider, a genuine price must still be paid. Insider privatization cannot be allowed to degenerate into a concealed form of giveaway.
- The third characteristic of the organic development strategy follows from the second: any giveaway of state property must be avoided.
- Preference must be given to sales that produce a dominant owner. This may be a businessperson, a group of owners, or a privately owned company (domestic or foreign) with a history of private ownership. A particularly desirable owner is a strategic investor who is prepared to inject significant capital into the company. Where the form of a public limited company is chosen, there is no need to avoid dispersing some of the shares. But wherever possible, every company should have a core owner.
- Companies must face harder budget constraints to ensure the financial discipline essential to a market economy. New laws will have to be passed—including for bankruptcy, accounting, and banking—and consistently enforced. The trinity of privatization, liberalization, and stabilization will not ensure a successful transition. Harder budget constraints are just as important.

State enterprises with chronic losses do not need to be privatized at all costs or sustained artificially for too long. As the budget constraint hardens, it performs a process of natural selection among them. Profitable enterprises can be sold sooner or later. Enterprises that cannot be sold, because they have zero or a negative value, must enter bankruptcy proceedings, not be given away. Privatization

through bankruptcy and liquidation is one of the main techniques for changing ownership.

The private sector's share of gross production will grow as new private businesses emerge and the state sector shrinks. The state shrinks in two ways: state enterprises are sold to private owners, or they go bankrupt and exit.

The Accelerated Privatization Strategy

The second strategy I call the strategy of accelerated privatization. It has three main features:

- The most important task is to eliminate state ownership as quickly as possible.
- The main privatization technique is some form of giveaway—such as a voucher scheme—that distributes freely and equally among the country's citizens the property rights to the state companies being privatized. This approach may tolerate or even encourage takeovers by managers. In many cases this turns out to be a pseudo management buyout because the managers pay a very low price, which is almost tantamount to receiving the property rights free of charge.
- There is no need to avoid dispersed ownership. In fact, it may be preferred. What needs to be emphasized is that all citizens will share in the property rights of former state enterprises so that “people's capitalism” develops.

This strategy has two fewer features than the strategy of organic development. Advocates of accelerated privatization also approved of bottom-up private sector development, but they did not emphasize it in their proposals—while it was placed at the fore of ownership reform by advocates of organic development. Similarly, supporters of accelerated privatization would have approved of hardening budget constraints in principle. They did not press for retaining soft budget constraints, but the requirement of a hard budget constraint got lost in their proposals, and not by chance. I will return to this in the context of the Czech and Russian experiences.

The most important differences between the two strategies are not the items in each group of features, but the items that receive greatest emphasis. That is, where do the strategies suggest focusing political attention, legislative and administrative capacity, intellectual interest, and research activity? The two strategies are very different in this respect. The organic development strategy emphasizes healthy growth of the new private sector, while the accelerated privatization strategy pushes for rapid liquidation of the state sector.

The Road to a Free Economy and other writings of mine that appeared at about the same time outlined and recommended the organic development strategy. I was not alone in doing so: quite a few other analysts offered similar views. I greatly appreciate the positions taken by Andreff (1992), Bolton and Roland (1992), Brabant (1992), McKinnon (1992), Murrell (1992a, b, and c), Murrell and Wang (1993), and Poznanski (1993). But supporters of organic development were limited

to a small minority of Western academic economists. Most of the profession supported the accelerated privatization strategy, often using aggressive arguments to do so. Today I am certain that the organic development strategy was correct. The accelerated privatization strategy was inferior at best and harmful at worst.³

Intellectual Underpinnings of the Strategies

What were the intellectual inspirations for the two strategies? No one offered a strict line of thinking or produced a model that drew conclusions from precisely formulated assumptions. Advocates of both strategies blended knowledge from economics with intuition or with some vision of how capitalism was going to develop. In rereading the writings of those times, my purpose was not to discover which authors are cited. It was more to read between the lines to determine which ideas inspired the visions. I realize that I am treading on uncertain ground and could offer an inaccurate interpretation. Nonetheless, I will try to answer the question.

Let me begin with the easier part of the task, introspection. Which works and intellectual strands influenced me the most as I thought about ownership reform at the end of the 1980s? One source was Friedrich Hayek—especially his ideas on the development of the market economy and his opposition to “constructivism” (Hayek 1960, 1989). I thought it grotesque that my Czech colleagues, while referring to Hayek on several occasions, should be concocting the rules of the game for the voucher scheme and state prescriptions for putting it in practice. Hayek attached enormous importance to the spontaneity of capitalism, to the way it picks out, by evolutionary means, viable institutions capable of survival.

Another of my intellectual sources was Joseph Schumpeter—not the Schumpeter of *Capitalism, Socialism, and Democracy* (1976 [1942]), who places naive hope in market socialism, but an earlier Schumpeter (1968 [1911]), who identifies the entrepreneur as the central figure of capitalism. Schumpeter’s market economy is not a sterile, equilibrium-bound, Walrasian world. It is a world of real rivalry, where people found new firms, conquer new markets, and introduce new products. I felt that Eastern Europe, after its numbing dose of bureaucracy, needed tens of thousands of Schumpeterian entrepreneurs. Closely connected with this is Schumpeter’s well-known idea of creative destruction. This combines in my current thinking with the hardening of budget constraints and the painful but essential natural selection that ensues. A powerful process of exit and entry is the driving force for reallocating resources from less to more productive firms (Caballero and Hammour in this volume).

A third source was the image of the beginnings, development, and consolidation of capitalism that formed in my mind from various readings. These included the French Annales school, the writings of Fernand Braudel and others (which clarify the evolutionary nature of the process, especially Braudel 1985 [1975]), and studies of the commercial laws and financial discipline introduced under early capitalism.

Finally, I was strongly influenced by the study of socialist systems. I did not use the word *institution* in every other paragraph as has recently become fashionable.

But I understood what a system is, and the difference between socialism and capitalism. And I was aware that this difference would not disappear just by privatization, stabilization, and liberalization.

What intellectual sources influenced the advocates of accelerated privatization, inspiring them to produce their vision of how to “construct” capitalism at a rapid pace? It is not sufficient to refer in general terms to the influence of mainstream economics. Even if the adherents of accelerated privatization do not refer to them, I am convinced that they were strongly influenced by two authors. One (by a twist of fate) was Karl Marx and the other was Ronald Coase. I concede that they make strange bedfellows.

Sophisticated Marxists would call the accelerated privatization strategy vulgar Marxism. I would add that what it took from Coase is vulgar Coaseism as well. Here vulgar Marxism means a simplified formula: the change in ownership is not just a necessary condition of capitalism, but a sufficient one. Capitalist property relations form the base that goes on to create its own superstructure: the institutions, political organization, and ideology required to operate the capitalist base.

History and the postsocialist transition show that the relationship between base and superstructure is far more complicated. The mere existence of capitalist property relations is insufficient for the consolidation of capitalism. Transformation of the economy and the society often proceeds in parallel, with many interactions. One sphere advances and then the other, reacting to the first. There is no universal rule governing the sequence of the interactions. If a drastic reform of ownership should precede the transformation of political, legal, and cultural institutions, the institutional transformation may follow slowly and painfully, at great social cost. So, even if it is feasible, it is not certain that having rapid and drastic ownership reform before the transformation of auxiliary institutions is the most beneficial sequence.

For the simplified formula of vulgar Coaseism, it does not matter if the initial allocation of legal entitlements is inefficient. An efficient allocation will ultimately appear so long as:

- The exchange occurs in a perfectly competitive market.
- The exchange is free, and there are no barriers to recontracting.
- The recontracting involves no or low transaction costs (Coase 1960).

But what if these conditions do not apply? In fact, that is often the case in transition economies: there are serious problems with these conditions. The renegotiation and recontracting of the allocation of property rights may be blocked by powerful interest groups (as in Russia). Moreover, appalling social costs can arise during the reallocation period.

Four Countries' Experiences

In transition economies there is a close causal relationship between healthy development of the private sector, hardening of the budget constraint, forceful restructuring of production, and as the ultimate result, growth of labor productivity. Here

growth of labor productivity is more expressive than growth of per capita GDP because it sheds a clearer light on the effect of restructuring. Socialism left a legacy of mass unemployment on the job. The organic development strategy seeks to dispose of this legacy even if it means taking painful and unpopular measures. The accelerated privatization strategy shrinks from doing so. In 1998 labor productivity in Hungary was 36 percent higher than in 1989, while in Poland it was 29 percent higher. In the Czech Republic it was just 6 percent higher than in the last year of socialism. And in Russia labor productivity was 33 percent lower in 1998 than in 1989 (Economic Commission for Europe 1999).

Hungary has pursued the organic development of its private sector, coming closest to exhibiting the five features of the strategy described above.⁴ Still, Hungary's experience should not be idealized. Many misuses occurred—which can arise not only with free distribution but also with privatization by sale. Although none of the great corruption scandals came to a head, experts and the public suspect that abuses were not rare.

Nonetheless, Hungary's economic achievements have been impressive. Hundreds of thousands of small and medium-size firms came into being. Hardening of the budget constraint in the first half of the 1990s allowed natural selection to sweep over corporations. This coincided with a perceptible strengthening of financial discipline. The chains of mutual debt among companies were broken, and private contracts improved. A start was made on consolidating banks. All these developments attracted considerable foreign capital. Such capital was one of the main factors explaining Hungary's increased productivity and exports.

Poland occasionally flirted with the accelerated privatization strategy, but economic policy remained close to the organic development strategy. Many Polish economists now recognize that, apart from macroeconomic stabilization, Poland's success stemmed from massive new entries, vigorous bottom-up growth of the private sector, and foreign capital inflows (Dabrowski, Gomulka, and Rostowski forthcoming).

In the early 1990s the leaders in what became the Czech Republic were the first who wanted to apply the accelerated privatization strategy. Václav Klaus, the country's economist prime minister, championed the voucher scheme, which was applied energetically.⁵

Many analysts have offered explanations for why the voucher program did not yield the expected results (see Coffee 1996, 1998; Ellerman 1998; Nellis 1999; and OECD 1998, 2000). In the first phase, state assets were dispersed among millions of voucher owners—only to be concentrated again afterward in investment funds. But the funds lacked the capital to develop the backward companies or put in real investment. The funds were intertwined with large commercial banks, which were mainly or entirely owned by the state. This ownership structure was incapable of building strong corporate governance. Restructuring dragged on. Despite the strident free enterprise rhetoric directed at the outside world, budget constraints remained soft. Whereas privatization by sale engenders natural selection, the transfer of property rights by giveaway maintains the existing structure.

The Czech Republic's performance has been disappointing. The accelerated privatization strategy seems to explain many of the problems, though serious mistakes in macroeconomic policy also contributed to the lagging economy.

Russia is perhaps the saddest example of the failure of the accelerated privatization strategy. Here every feature of the strategy appeared in an extreme form: a voucher scheme was imposed, coupled with mass manipulated transfers of property into the hands of managers and privileged bureaucrats. In this environment an unprecedented ownership reform occurred—in which the ownership of natural resources, especially oil and gas, was expropriated by oligarchs.⁶

Russia's woes are closely linked to the survival of the soft budget constraint, in a form that has infiltrated and damaged every cell of the economy and body politic. Russia has become a "nonpayment society" (Pinto, Derbentsov, and Morosov 1999). Companies do not pay their suppliers, employers do not pay their employees, and debtors do not pay their lending banks. All this is tolerated by the executive and the judiciary. In fact, the state often sets a bad example by falling behind on the wages and insurance contributions of state employees and on pensions.

A Review of Four Arguments

Bearing in mind the features, ideas behind, and outcomes of the two strategies for ownership reform and private sector development in transition economies, consider some of the arguments heard in the debates of the early 1990s. First, advocates of accelerated privatization eagerly cited ethical considerations. It was argued that, for reasons of fairness, every citizen had to be given an equal share of state property. Experience has shown that this is a hypocritical argument. The initial allocation of property remained for only a short period, then gave way to a highly concentrated ownership of former state property. In Russia this led to an absurd, perverse, and extremely unfair form of oligarchic capitalism.⁷

If the sale of state assets occurs at a correct price, it does not alter the distribution of wealth or income. The state's wealth is not reduced; it simply changes form. Revenue from privatization has to be invested, not consumed. Hungary used its receipts to reduce foreign debt, at least during the big wave of privatization, when much of the energy and telecommunications sectors were sold. The consequent reduction in interest payments and improvement in the country's credit rating brought real benefits for all Hungarians.

Second, advocates of organic development placed great emphasis on sociological issues. The bourgeoisification of society, with the development of a property-owning class, is essential to the consolidation of capitalism. At a certain stage in the maturity of capitalism, a great role is played by the fragmented ownership of shares, coupled with institutional ownership. But there can be no running ahead. The appearance of institutional investors cannot substitute for a transformation in the stratification of society. This argument was confirmed by the first decade following socialism. There is a close correlation between measures of economic success and the restratification of society.

Third, the arguments that most appealed to economists involved economic efficiency. Here the organic development strategy is far superior. New private companies are generally more productive than those that remain under state ownership or those that were privatized during the transition (Konings, Lehmann, and Schaffer 1996; Konings 1997). The Schumpeterian spirit of enterprise—sweeping aside inefficient, nonviable companies; enabling new, real owners intent on establishing order; attracting foreign capital glad to make large, modern investments—boosts productivity growth and enhances export performance.

Finally, there are the political arguments. The voucher program was crucial to the victory of the governing party in the second free Czech elections—the only case in Eastern Europe in the past 10 years where the same government continued for a second term. By that yardstick, privatization was a success.⁸ By contrast, the coalitions that ruled in the first parliamentary cycles in Hungary and Poland failed in the second general elections. The rival coalitions that took office pursued much the same organic development strategies as their predecessors. Four years later, after abstaining from using giveaway privatizations as an election weapon, they suffered defeat. So accelerated privatization proved more favorable according to the Machiavellian criterion of retaining power.

Advocates of accelerated privatization everywhere, especially in Russia, repeatedly argued that if the window of opportunity opens for privatization, the opportunity has to be seized and the privatization carried out rapidly. It has to be done while the state bureaucracy is still confused, weak, and unable to resist. And the change in ownership has to be made irreversible, lest there never be another chance of doing so.

This argument can be neither confirmed nor denied by purely logical, speculative means. No argument can be sufficiently defended. Although it is clear today that Czech democracy, for instance, was not under any threat of a communist restoration or a reappearance of Soviet tanks, that was not so clear in 1991.

Reassessing events in Russia is especially problematic from this point of view. It has been repeatedly argued that mass privatization had to be carried out swiftly before the Communist Party gained its electoral victory. No type of privatization could have been pushed through the Duma once the Communist Party had become the tone-setter there.

I think that a faulty, upside-down causal explanation lies behind this argument. If privatization had taken another course, one without so many glaring abuses and social losses, there would not be such strong nostalgia in Russia for the communist system. An ownership reform thrust on society may be irreversible. But a more solid foundation for an irreversible advance of capitalism would be established if a broad bourgeoisie developed, property rights and private contracts were applied consistently, democracy were institutionalized, and the market economy enjoyed political support from most voters.

Macroeconomic Stability

In reviewing *The Road to a Free Economy*, I felt satisfied as I read the chapter on privatization. I cannot say the same about the chapter on stabilization. If some

miraculous time machine could take me back to that time (with my thoughts as they are today), I would rewrite the stabilization chapter before sending the book to the printer. The chapter dealt with several questions, three of which I address here.

Timing

When I wrote the book in 1989, Hungary was suffering from severe macroeconomic problems that required strong correction. It was clear that the adjustment would be painful, begging the question of when it should be done. My book recommended doing it right away—in the next year or two. That recommendation was soon repeated in the Czech, Polish, and Russian editions of the book and in several other foreign editions.

The main argument was that a new chapter of history was being opened. At that moment a freely elected government would have the moral legitimacy to ask the public to make a sacrifice. It would still be possible to claim that the government was trying to remedy the previous regime's worst omissions (as well as those that could be corrected most swiftly). If adjustment were postponed, people would feel, justifiably or unjustifiably, that the troubles had been caused by the democratically elected government, not the previous system.

I still believe that this position is correct. A dramatic step of this kind was taken in Poland in 1990 with Finance Minister Leszek Balcerowicz's reform program. And while I criticized the Czech government earlier in this article, I would like to pay tribute to the Klaus government for the bold adjustments it made to macroeconomic policy in 1991.

As a Hungarian, I sincerely regret that the government of my country rejected swift macroeconomic adjustment and the opposition did not press for it. Those decisions depended on political will, not objective economic conditions. Leading politicians were afraid to take unpopular action. As a result adjustment was postponed through the first four-year Parliament until eight months into the second. It was eventually adopted in 1995, when Hungary came close to financial collapse in the wake of Mexico's peso crisis. Careful advice was not enough. It took catastrophic signals at the frantic last minute before the government could bring itself to take corrective measures to avert the crisis.

Most experts agree that this postponed adjustment cost more than it would have had it been implemented earlier. Such a decision is not made in a purely rational, economic context. There is an ethical and political dilemma. It is a question of the intertemporal distribution of pain and gain, and of acceptance of the political price of unpopular measures.

Predictions

The proposals I made in *The Road to a Free Economy* rested on my forecasts of the macroeconomic consequences of the transition. But my prognosis was wrong. I did not predict the deep recession that followed; I was too optimistic in my expectations

of future growth. Many of my colleagues in Hungary and abroad made more realistic predictions.

I fault myself because I had information on which I could have based better forecasts. For instance, I could have read more carefully my own forthcoming book, *The Socialist System* (Kornai 1992b), which might have initiated in me the following line of thinking:

- The socialist system left behind a badly distorted structure of input and output. Correcting this called for creative destruction. But while destruction is rapid, creation is much slower. The balance of the two processes implied that there would be a deep recession.
- The socialist system established a special mechanism for coordinating activities. Although this mechanism was inefficient, it did operate. With the change of system, the old mechanism broke down, but the new market mechanism had not yet managed to take over all the tasks of coordination. In a study I wrote later on the transformational recession (Kornai 1993b), I called this situation an institutional no-man's land.

These changes, along with other factors, led to the region suffering the deepest recession in international economic history. The classic recipes for macroeconomic stabilization had to be altered and augmented before any adjustment and transformation could succeed.

What Action at One Stroke Can Achieve

The Road to a Free Economy recommended that a radical program of action be taken at one stroke. Even today I do not reject the notion of a radical adjustment package in which several measures are taken simultaneously. A well-compiled package of properly calibrated measures can restore equilibrium in several important dimensions of the macroeconomy all at once, or at least bring the economy much closer to a tolerable degree of disequilibrium (for instance, reducing the current account deficit or the budget deficit to a sustainable level).

What I criticize in that proposal today is its misplaced emphasis. Too much attention was paid to what could be achieved rapidly with a drastic adjustment package—and too little to how to consolidate this quick fix and produce a further, lasting improvement.

It is hard to achieve economic equilibrium, but very easy to lose it again. During the 1990s it seemed time and again in the Czech Republic, Hungary, Poland, and Russia as if the macroeconomy was on the right track. Then came another jolt: deceleration or even deterioration in certain indicators. For growth to be sustainable, there has to be not just one macroeconomic intervention, but deep, comprehensive institutional reforms.

My book dealt adequately with everything connected, directly or indirectly, to the budget constraint. But I cannot acquit myself of a mistake also made by many others: not pointing sufficiently to the importance of other reforms. It is easy to rapidly improve the budget balance at a single stroke by, say, raising rates on existing taxes.

But a lasting improvement requires radical tax reforms, a broader tax base, the introduction of new taxes, and consistent tax collection. And that is only one side of fiscal reform, perhaps the easier side. The other side requires cutting state spending by reorganizing the state apparatus and the financing of education, health care, and other welfare systems. Similarly, it is relatively easy to declare that the currency is convertible. It is much harder to organize an effective system of international payments, to develop strong connections between domestic and international banking systems, and to guarantee that international payment agreements will be observed.

Macroeconomic stabilization is not a battle, but an endless war. Stabilization cannot be gained by a blitzkrieg. Institutional reforms can only be obtained step by step, through a series of larger and smaller reforms. I regret that this idea did not feature in *The Road to a Free Economy*.

Conclusion

The debates of the early 1990s involved choosing between gradualism and shock therapy. But I believe that was a false dichotomy, so I will not try to resolve it. It implies a single yardstick: speed. While speed is important, it is not the primary measure of success. In those days many participants in the postsocialist transformation were obsessed with speed.⁹ The Czech Republic was congratulated on being the first country to privatize the bulk of its economy. Hungary's private sector was not making an equivalent contribution until two years later, and Poland's not until perhaps three years later. But so what? The transformation of society is not a horse race. The main indicator of success is not who passes the winning post first.

Excessive emphasis on speed leads to impatience, aggressiveness, and arrogance. Ironically, the expression mass privatization, used as a synonym for giveaways and voucher schemes, is the inverse of the mass collectivization familiar under Stalin. Stalin did not want to spend much time bothering with voluntary collectivization. Using merciless violence, he imposed collective ownership on the peasantry within two to three years. I do not want to exaggerate the comparison—no gulags or brutality were required in the 1990s. The forcing of change was milder. Still, there were similarities: the subordination of ownership reform to political purposes, the horror felt toward gradual change, the impatience, and the obsession with speed.

The transition from socialism to capitalism has to be an organic development. It is a curious amalgam of revolution and evolution. It is a trial and error process that retains or liquidates old institutions and tries out, accepts, or rejects new ones. Each element in the process might be very rapid, fairly rapid, or slow. Each has its own appropriate speed. Some episodes call for one-stroke intervention. Many other processes advance through incremental changes.

There are more important criteria than speed. I start from the conviction that capitalism is superior to socialism. From that it follows that the stronger are capitalism's foundations, the better will be the system's medium- and long-term performance. So the emphasis must be placed on consolidation and stability—and at the same time, on making growth sustainable, not on breaking records with it.

Notes

1. This article mainly deals with Kornai (1990), but a few other lectures and publications at the start of the transition allowed me to clarify my views. The Tinbergen lecture (Kornai 1992a), delivered in 1991, concerned privatization. The Myrdal lecture (Kornai 1993a), delivered in 1992, was about hardening budget constraints. Both lectures are assessed as part of this evaluation.

2. In a narrow sense the word *adviser* means a person who is invited to advise a government, state, international organization, or political party or movement. Many economists inside and outside transition economies did so at the beginning of the transition. I turned down all such invitations. But there is a broader meaning to the word adviser: a person who does research and makes policy recommendations without being commissioned to do so. As the author of *The Road to a Free Economy*, I consider myself an adviser in the broader sense. When I was young, just before the 1956 Hungarian revolution, I belonged to a group that recommended reforms. After the defeat of the revolution, 33 years passed in which I never again drew up another comprehensive economic policy proposal. I concentrated on research. My role did not change radically until the first free elections were announced in Hungary, when I realized that it was time to present proposals that had formed in my mind.

3. Dyck (in this volume) shows that most transition economies using direct sales and concentrated ownership with openness to outsiders had growth rates higher than the mean for the region. Countries adopting voucher schemes with dispersed ownership had growth rates lower than the mean.

4. It is not possible to say how much *The Road to a Free Economy* influenced the Hungarian administrations that succeeded each other at four-year intervals. Politicians do not usually acknowledge their intellectual debts. At the time the book was hotly debated in Hungary—not only in the specialist press but also in daily newspapers and on radio and television. Many leading politicians and their advisers must have read it.

5. The idea of vouchers did not originate in the Czech Republic. It had appeared earlier in Poland, in a paper by Lewandowski and Szomburg (1989). Of the Czech program, Klaus (1997, p. 72) wrote in 1992, “Our nonstandard voucher privatization proved to be rapid and efficient.”

6. For a critical analysis of the consequences of Russian privatization, see Black, Kraakman, and Tarassova (2000). About the barriers of free entry, see Broadman (2000) and Desai and Goldberg (2000).

7. I am not claiming that privatization by sale, as opposed to giveaway, is necessarily “clean.” I mentioned earlier that there were presumably several shady transactions in Hungary. All I seek to do here is to refute the claim that free distribution, by its nature, is fair. In any case, most Russians looked on vouchers with suspicion from the outset and did not expect them to appreciably improve their financial position (see Blasi, Kroumova, and Kruse 1997, pp. 76–77).

8. Though the same government fell two years later in the middle of the parliamentary cycle—not least because of the economic policy mistakes it had made.

9. Anatoly Chubais, the leading figure in Russian privatization, gave a lecture to the Carnegie Endowment for International Peace in May 1999. “Asked about his role as privatization minister from 1992 to 1994, Chubais conceded that his privatization efforts could be characterized as ‘Bolshevik-style’—lacking public support and quickly executed . . . His strategy was to privatize as quickly as possible, using every minute of the day to privatize: ‘I did not speak, I privatized,’ Chubais proclaimed” (CEIP 1999).

References

- Andreff, Wladimir. 1992. “French Privatization Techniques and Experience: A Model for Central-Eastern Europe?” In Ferdinando Targetti, ed., *Privatization in Europe: West and East Experiences*. Aldershot, U.K.: Dartmouth.

- Black, Bernard S., Reinier Kraakman, and Anna Tarassova. 2000. "Russian Privatization and Corporate Governance: What Went Wrong?" *Stanford Law Review* 52 (6): 1731–1808.
- Blasi, Joseph R., Maya Kroumova, and Douglas Kruse. 1997. *Kremlin Capitalism: The Privatization of The Russian Economy*. Ithaca, N.Y.: Cornell University Press.
- Bolton, Patrick, and Gerard Roland. 1992. "Privatization Policies in Central and Eastern Europe." *Economic Policy* 15 (October): 275–309.
- Brabant, Jozef M. van. 1992. "Divestment of State Capital." In Kazimierz Z. Poznanski, ed., *Constructing Capitalism: The Reemergence of Civil Society and Liberal Economy in the Post-Communist World*. Boulder, Colo.: Westview Press.
- Braudel, Fernand. 1985 [1975]. *The Wheels of Commerce*. London: Fontana Press, Imprint of HarperCollins.
- Broadman, Harry G. 2000. "Reducing Structural Dominance and Entry Barriers in Russian Industry." World Bank, Europe and Central Asia Department, Poverty Reduction and Economic Management Unit, Washington D.C.
- CEIP (Carnegie Endowment for International Peace). 1999. *Meeting Report, May 1999*. Washington, D.C.
- Coase, Ronald. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 17 (2): 357–76.
- Coffee, John C. Jr. 1996. "Institutional Investors in Transitional Economies: Lessons from the Czech Experience." In Roman Frydman, Cheryl Gray, and Andrzej Rapaczynski, eds., *Corporate Governance in Central Europe and Russia*, vol. 1. New York: Central University Press.
- . 1998. "Investing in a Corporate Monitor for Transitional Economies: The Uncertain Lessons from the Czech and Polish Experiences." In Klaus J. Hopt, Hideki Kanda, Mark J. Roe, Eddy Wymeersch, and Stefan Prigge, eds., *Comparative Corporate Governance*. Oxford: Clarendon Press.
- Dabrowski, Marek, Stanislaw Gomulka, and Jacek Rostowski. Forthcoming. *Whence Reform? A Critique of Stiglitz Perspective*. Warsaw: Centre for Social and Economic Research.
- Desai, Raj M., and Itzhak Goldberg. 2000. "The Vicious Circles of Control: Regional Governments and Insiders in Privatized Russian Enterprises." Policy Research Working Paper 2287. World Bank, Washington, D.C.
- Economic Commission for Europe. 1999. *Economic Survey of Europe 1999 No. 3*. New York: United Nations.
- Ellerman, David. 1998. "Voucher Privatization with Investment Funds: An Institutional Analysis." Policy Research Working Paper 1924. World Bank, Washington, D.C.
- Hayek, Friedrich A. 1960. *The Constitution of Liberty*. Chicago: Chicago University Press.
- . 1989. *Order—With or Without Design?* London: Centre for Research into Communist Economies.
- Klaus, Václav. 1997. *Renaissance: The Rebirth of Liberty in the Heart of Europe*. Washington, D.C.: Cato Institute.
- Konings, Jozef. 1997. "Firm Growth and Ownership in Transition Countries." *Economic Letters* 55: 413–18.
- Konings, Jozef, Hartmut Lehmann, and Mark E. Schaffer. 1996. "Job Creation and Job Destruction in Transition Economy: Ownership, Firm Size and Gross Job Flows in Polish Manufacturing 1988–91." *Labor Economics* 3 (3): 299–317.
- Kornai, János. 1990. *The Road to a Free Economy: Shifting from a Socialist System—The Example of Hungary*. New York: W. W. Norton. [Also available in Hungarian: 1989. *Indulatós röpírat a gazdasági átmenet ügyében*. Budapest: HVG Kiadó.]

- . 1992a. “The Principles of Privatization in Eastern Europe.” *De Economist* 140 (2): 153–76.
- . 1992b. *The Socialist System. The Political Economy of Communism*. Princeton, N.J.: Princeton University Press.
- . 1993a. “The Evolution of Financial Discipline under the Postsocialist System.” *Kyklos* 46 (3): 315–36.
- . 1993b. “Transformational Recession: A General Phenomenon Examined through the Example of Hungary’s Development.” *Economie Appliquée* 46 (2): 181–227.
- Lewandowski, Janusz, and Jan Szomburg. 1989. “Property Reform as a Basis for Social and Economic Reform.” *Communist Economies* 1 (3): 257–68.
- McKinnon, Ronald. 1992. *The Order of Economic Liberalization: Financial Control in the Transition to a Market Economy*. Baltimore, Md.: Johns Hopkins University Press.
- Murrell, Peter. 1992a. “Conservative Political Philosophy and the Strategy of Economic Reform.” *East European Politics and Societies* 6 (1): 3–16.
- . 1992b. “Evolution in Economics and in the Economic Reform of Centrally Planned Economies.” In Christopher Clague and Gordon C. Rauser, eds., *The Emergence of Market Economies in Eastern Europe*. Oxford: Blackwell.
- . 1992c. “Evolutionary and Radical Approaches to Economic Reform.” *Economics of Planning* 25 (1): 79–95.
- Murrell, Peter, and Yijiang Wang. 1993. “When Privatization Should Be Delayed: The Effect of Communist Legacies on Organizational and Institutional Reform.” *Journal of Comparative Economics* 17 (2): 385–406.
- Nellis, John. 1999. “Time to Rethink Privatization in Transition Economies?” *Finance and Development* 36 (2): 16–19.
- OECD (Organisation for Economic Co-operation and Development). 1998. *Economic Survey of the Czech Republic, 1998*. Paris.
- . 2000. *Economic Survey of the Czech Republic, 2000*. Paris.
- Pinto, Brian, Vladimir Derbentsov, and Alexander Morozov. 1999. “Dismantling Russia’s Non-payments System: Creating Conditions for Growth.” World Bank, Europe and Central Asia Department, Poverty Reduction and Economic Management Unit, Washington, D.C.
- Poznanski, Kazimierz Z. 1993. “Poland’s Transition to Capitalism: Shock without Therapy.” In K. Z. Poznanski, ed., *Stabilization and Privatization in Poland*. Boston: Kluwer Academic Publishers.
- Roland, Gérard. 2000. *Transition and Economics: Politics, Markets, and Firms*. Cambridge, Mass.: MIT Press.
- Schumpeter, Joseph A. 1968 [1911]. *The Theory of Economic Development. An Inquiry into Profits, Capital, Credit, Interest and Business Cycles*. Cambridge, Mass.: Harvard University Press.
- . 1976 [1942]. *Capitalism, Socialism, and Democracy*. New York: Harper and Row.

*New Development
Thinking*

Consensus Building, Knowledge, and Conditionality

Paul Collier

Progress in poverty reduction depends primarily on policy and institutional changes in low-income countries. The World Bank's previous approach to inducing these changes relied on negotiated conditions on loans, or conditionality. The empirical evidence suggests that this approach was largely ineffective—where change occurred, it was chosen by governments rather than induced by conditions on loans. Various countervailing pressures undermine the effectiveness of loan conditionality. An alternative approach to inducing change is to empower, through knowledge and participation, domestic constituencies to make change. This approach is likely to be more effective in promoting policy change—and essential in promoting institutional change, now usually the frontier of economic reform. This shift in focus is part of the rationale for the Bank's Comprehensive Development Framework.

Our understanding of why some countries stay poor and what donors can do to reduce poverty has evolved over the past 30 years. I start by looking briefly at how our ideas about why countries stay poor have changed, but focus primarily on how our ideas about the role of the World Bank and of donors have changed.

Why Poverty Persists

Thirty years ago the diagnosis of persistent mass poverty was pessimistic. People feared that the poor would stay poor even if growth could be achieved. The evidence for the Kuznets curve appeared to suggest that growth in poor countries was intrinsically biased against the poor: as low-income countries grew richer, inequality increased. In this view poverty reduction would require active, large-scale redistribution policies to offset the forces that seemed to systematically exclude the poor from the benefits of growth.

Paul Collier is director of the Development Research Group at the World Bank.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

In contrast to the assumption about mass poverty, the diagnosis of slow growth was highly optimistic. Slow growth was initially seen as a result of low savings, which in turn were seen as a result of low incomes. Sufficiently large aid flows, it was argued, would spring a society from this savings trap. Analysts soon realized that low savings could not be the primary cause of slow growth because in much of the developing world the capital stock was much lower than savings rates indicated. For example, by 1990 about 40 percent of the private wealth of Africa and the Middle East had shifted offshore (Collier, Hoeffler, and Pattillo 2001). These regions evidently were short of capital domestically at least in part because they had hostile environments for economic activity. The diagnosis of slow growth therefore changed from the savings trap explanation to an explanation of what made these environments hostile to investment, but it remained optimistic. The culprits were a shortlist of dysfunctional economic policies—high inflation, trade barriers, overvalued exchange rates, and restrictions on private activity—all of them readily fixable.

We now think that these positions on poverty and growth were wrong.

The Effect of Growth—and Pro-Growth Policies—on Poverty

Our analysis of mass poverty was too pessimistic. Although growth does not necessarily help the poor, it does not necessarily exclude them either. With reasonable care, we can make growth sufficiently broad based to eliminate *mass* poverty (Li, Squire, and Zou 1998; Dollar and Kraay 2000). Pro-growth policies have received an undeservedly hostile response from the development studies community. This response must be contested because it undermines, and is intended to undermine, the political basis for growth.

Take the example of Uganda, clearly the testing ground for pro-growth policies in the 1990s. A recent so-called “participatory assessment” of how poverty had changed in Uganda, using the standard methods of development studies, concluded that “the rich are getting richer and the poor are getting poorer” (Uganda 1999, p. 27). This message would normally have been used to discredit the pro-growth policies that the Ugandan government had pursued so effectively. Was it correct?

Fortunately, in this case the participatory assessment results could be challenged by objective evidence. Large, national, random sample consumption surveys had been conducted annually throughout the decade. The results of these surveys showed precisely the opposite of what the assessment had concluded: consumption poverty had fallen in every year of the survey, and the cumulative reduction in poverty was very substantial (Appleton 2001). Moreover, inequality had been reduced. The poor had benefited disproportionately from Uganda’s growth, although 80 percent of the reduction in poverty came from growth rather than redistribution.

If we broaden the definition of poverty to include not only consumption but also education and health services, the evidence is even more positive. During the 1990s school enrollment increased dramatically and infant mortality declined sharply. If we go further and include the notion of empowerment in our definition of poverty,

we find that during the 1990s the Ugandan population acquired meaningful electoral power, gained access to competing sources of information through a free press and radio, and was liberated from monopoly crop marketing. Thus the claim that the rich were getting richer and the poor were getting poorer grossly mischaracterized the Ugandan experience.

How could the participatory assessment method produce such an erroneous finding? The authors of the report breezily mention in a footnote the sharp fall in consumption poverty and offer no explanation of the contradiction between this finding and their own results. The author of a later attempt to defend the results of the assessment is forced to admit that data collection was politicized—some people were paid to participate—and that the improved environment in Uganda has raised expectations and empowered people to complain more about unsatisfactory social services than they might have before (McGee 2000). While the author reassures us that the participatory method deals with these problems in a holistic way, some might conclude that a method that represents an improvement as a worsening, without even acknowledging this severe limitation, is suspect.

The theory of fads and information cascades gives us an insight into why the participatory assessment method is liable to produce such misleading results (Bikhchandani, Hirschleifer, and Welch 1998). False opinions can become dominant when people assume that the collective opinions of others contain more information than their own limited experience: the idea that growth leads to worsening poverty becomes accepted through a cascade of gossip. In effect, the participatory assessment method allows development practitioners to hold up a mirror to their own prejudices while interpreting the results as if they were hard data.

Of course, pro-growth policies can be more or less effective in reducing poverty. Recent research in the World Bank has explored the impact of various pro-growth policies on the poor. Dollar and Kraay (2000) show, for example, that low inflation disproportionately benefits the poor relative to other groups in society. This implies that a macroeconomic policy package that seeks to reduce poverty—rather than simply to promote growth—should give more weight to preventing inflation.

While growth will eliminate *mass* poverty, and well-designed growth policies will eliminate it quite rapidly, growth will not be enough for some groups. That is why all developed societies have evolved sophisticated public welfare, health, and pension programs.

The design of redistribution systems, and especially of quick-acting social safety nets, in low- and middle-income countries requires a great deal of skill. The modernization of society inevitably weakens traditional family-based systems. There is also some evidence that the introduction of public safety nets further undermines traditional systems (Cox and Jimenez 1995). These are not arguments for continuing to depend on family-based assistance, but indications that a rapid transition to public systems may be necessary because traditional systems and public systems do not cohabit successfully. The rationale for combining policies to build these transfer systems with a pro-growth strategy is not to mitigate the ill effects of pro-growth policies, but to make these transfer systems more affordable.

Constraints on Growth

While our analysis of poverty was too pessimistic, our analysis of growth was too optimistic. The constraints on growth are not just a matter of macroeconomic policies. We increasingly recognize that while open market economies outperform closed, planned economies, markets need institutions and infrastructure.

Institutions matter. Surveys of African manufacturing show that firms across the continent are reluctant to do business with new clients because they lack effective means of contract enforcement. This reluctance evidently reduces competition and dynamism: firms are locked into their existing client base. Infrastructure matters too. A study in Uganda found that the most important constraint on private investment was unreliable electricity supply through the grid (Reinikka and Svensson 1999). Another study found that deficiencies in the telephone system were more detrimental to African growth than poor macroeconomic policies were (Easterly and Levine 1997).

Once gross macroeconomic misalignments are corrected, growth is generally constrained by the weakest point in a long list of factors. But governments cannot tackle all potential problems at once. Efforts across a broad front spread limited government implementation capacity too thinly, and priorities must be chosen. In the integrated development projects of the 1970s development agencies recognized that many factors might constrain development and tried to work on everything at once. These projects failed. Even when development agencies have been willing to prioritize, their choices of priorities have often reflected the latest fads in development thinking. For a time the most important constraint on growth was considered to be lack of openness, then it was lack of education, and then lack of transport infrastructure.

We need a diagnostic procedure that steers us toward more realistic conclusions. Every country must first work through an exhaustive checklist of the factors that might be the binding constraint on growth. Does the power work? Do the phones work? Do the courts work? Does the financial system work? Does health care work? Once the country determines the important constraints on growth, it should do a critical path analysis of efficient sequencing. If the courts don't work the banks won't work, because assets will not function as collateral. Until recently courts in Uganda would not transfer to banks control over assets pledged as collateral by defaulting borrowers. Defaulters could persuade the courts to delay decisions indefinitely. Not surprisingly, a survey of Ugandan banks found that their most pressing need from the government was a fast-track court procedure for collateral (Atingi-Ego and Kasekende 1997). Until the courts improve, Ugandan firms will inevitably be short of credit.

Another example of sequencing: if public sector employment reflects cronyism, staff will not be motivated to deliver services. Among Ghanaian public sector workers wages are 25 percent higher for those belonging to the locally dominant tribe, yet wages are unrelated to workers' numeracy and literacy skills (Collier and Garg 1999). Until hiring and promotions are skill based, the Ghanaian public sector will be unable to deliver such services as education and health care efficiently. A third example: if markets are not competitive, privatization will not work. In the Russian

Federation privatization has produced rapacious monopoly and in the process undermined the political constituency for private-led development.

The binding constraints on growth are not only country specific. Constraints evolve as policy interventions succeed: when one problem is fixed, another replaces it as the binding constraint. Thus governments need to update the entire checklist regularly.

Both governments and donors should understand the evolving constraints on growth. The ideal diagnostic procedure would therefore be collaborative.

How Donors Have Tried to Reduce Poverty

Coordinated action based on a common diagnosis would be a major change in the aid relationship, more likely to achieve donor objectives than previous approaches. Until now most donors have attempted to reduce poverty through project aid and aid for reform.

Project Aid

Thirty years ago donors tried to reduce poverty by delivering projects with a high rate of return. This approach encountered three problems. First, a project might be good in itself but not replicable, and a project that would not “scale up” was irrelevant in the larger context of economic growth. Second, projects were fungible: donor support for a project that a government would otherwise have financed itself freed resources for the government to use in other ways. In reality, the donor financed not the project it appeared to pay for, but the marginal project the government chose to undertake. Donor care in selecting among intramarginal projects thus made no difference to the overall portfolio of implemented projects (Feyzioglu, Swaroop, and Zhu 1997). Third, a project’s success or failure depended less on its design than on the environment in which it was implemented. Just like other investments, donor-funded projects failed in hostile environments (Isham and Kaufmann 1999).

Aid for Reform

Because good project design was not enough to achieve donor objectives of poverty reduction, in the 1980s donors developed a new approach—conditionality. They provided aid in return for explicit negotiated commitments to policy reform. The theory underlying this approach was that aid could be an incentive for policy change. But this theory implied that governments would undertake policy change that went against what they considered their interests, except for the receipt of aid. Policy change was the price governments would have to pay for aid. Put another way, donors would buy policy change with their aid.

The implications of using aid to induce reform were uncomfortable: if donors “bought” the reforms, they clearly “owned” them. When one African head of state became sufficiently annoyed with donors for complaining about the lack of political

rights in his country, he threatened to reverse the reforms unless they stopped complaining. The threat of reversing a reform program was intelligible only if both parties understood that the reforms belonged to the donors and not to the government (Collier 1997).

The implications of using aid as an incentive were so uncomfortable that donors came up with an alternative “fig leaf” theory: the notion of costs of adjustment. This theory held that policy change, because it was initially costly, was like an investment. Aid could finance the up-front costs.

Neither of these theories was completely wrong—sometimes the incentive of aid was enough to induce governments to implement policies they otherwise would not have undertaken, and sometimes such policy changes had up-front costs—but as general propositions both were dysfunctional. The lure of aid led governments to promise more than they intended to deliver and to implement more than they could sustain. For example, the government of Kenya sold the same agricultural reform to the World Bank five times in 15 years. The same condition has appeared in seven of the past eight policy framework papers prepared by the Bank for Malawi. The cumulative effect of such behavior was to destroy the credibility of governments, not only with donors but also with private investors.

Why Aid for Reform Has Not Induced Policy and Institutional Change

Adjustment costs are largely mythical. Most reforms, if they are sensible, lead to a rapid improvement in the economy. The emphasis on adjustment costs encouraged governments to exaggerate the difficulties of policy reform. The negotiating framework allowed rational governments to exaggerate the costs of policy change to maximize its price. Moreover, since donor negotiating teams saw their role as extracting the maximum reform for a given amount of aid, governments were always reluctant reformers at the margin, refusing to implement reform urged on them by donors.

Aid has not, on average, speeded policy change. A study of 220 reform programs shows that the success of such programs is systematically related to domestic political economy factors, such as how long the government has been in power, but is unrelated to donor behavior (Dollar and Svensson 2000). There is no overall relationship between aid flows and policy change. This is not as surprising as it might seem. While the incentives argument is elementary economics—the offer of aid for reform should induce a supply response—it ignores four offsetting effects.

Pressures Undermining the Effectiveness of Loan Conditionality

First, because aid alleviates governments’ fiscal and payments crises, it reduces the urgency of policy change. Economists would describe this as the income effect of aid offsetting the substitution effect. A recent study in collaboration with a team of African economists looked closely at 10 reform programs to see how the aid relationship affected the propensity to reform at different stages (Devarajan, Dollar, and

Holmgren 2001). In no case was sustained reform initiated by aid-for-reform packages. Reform was sustained only when initiated by governments, often in response to a crisis. In poor policy environments aid for reform thus paradoxically tends to delay reform—the income effect dominates the substitution effect.

Once reform was seriously under way, conditional aid was useful, although the key conditions were those chosen by governments. The usefulness of the aid resulted not from the substitution effect, but from governments' ability to signal priorities to their own bureaucracies by committing to two or three vital changes. Long lists of conditions diluted such signals and so were dysfunctional. As the reform process proceeded, the reforms became more complex and needed support from a wider constituency to be implemented effectively. By this stage conditionality served no purpose because the substitution effect could not be translated into effective incentives for such a large group of actors.

Second, aid for reform faces what economists call a time-consistency problem. If the government does not want the reform, it has little incentive to maintain it once the aid has been delivered. If it does want the reform, however, it normally does not need the aid to carry it out. The on-off pattern of Kenyan reforms noted earlier is an example of this time-consistency problem.

Third, aid for reform faces a moral hazard problem. The agencies that should enforce the conditions of aid for reform also happen to have an interest in seeing their loans repaid. Enforcing the conditions will reduce the probability of repayment. Svensson (1999) shows that loan disbursement has depended more on indebtedness than on adherence to conditions. Moral hazard has mattered.

Finally, while development agencies might have wanted to use conditionality to induce policy reform, other OECD interest groups with their own agendas tried to subvert conditionality to induce different behavioral change. Thaker (1999) claims to show statistically that loan approval by the board of the International Monetary Fund (IMF) in the early 1990s was significantly associated with a country's voting record at the United Nations. The more a country's voting pattern shifted toward the U.S. position, and the closer its voting was to the U.S. position, the higher the probability of loan approval. Although Thaker interprets this association as causal, other interpretations are equally possible. In any case we should not dismiss out of hand the possibility that political objectives subverted the economic objectives of conditionality.

Thus while aid for reform seems at first to be a straightforward application of economic incentives to promote behavioral change, such aid has encountered major obstacles in practice. The income effect has offset the substitution effect. Even when the incentives have been effective, they have induced policy oscillation rather than sustained reform. The agencies tasked with enforcing the bargains have had an incentive not to do so, and other interests have tried to hijack the bargaining.

Damage to Government Credibility

Although ineffective in achieving sustained policy improvement, aid for reform was effective in undermining the credibility of governments. Bad governments destroyed

their reputations by renegeing on the spirit of agreements, but even good governments faced a problem: when governments implemented and sustained policy reform, donors claimed the credit. This claim had some credibility because even reform-minded governments visibly resisted reform at the margin, while donors visibly pushed it.

Thus even governments that implemented reforms because they believed in them found it difficult to establish their claim to ownership with the investor community. The only policies governments were seen as truly owning were those that failed, because no one else claimed them. In the language of economics, governments need to be able to signal their true intentions to investors. Aid for reform made this signaling more difficult.

Misallocation of Aid

Aid for reform also diverted aid from countries where it could have been most effective in reducing poverty. We now know that poverty-efficient aid allocation takes into account three factors: the level of poverty, the level of policy, and diminishing returns. The level of poverty is straightforward: the more severe the poverty, the more effective aid is in reducing it. Thus for a given level of policy and institutions, the greater the poverty, the larger the aid program should be.

The second factor, the level of policy and institutional performance, is also fairly straightforward: the better the policy and institutional environment, the more effective aid is in increasing growth and reducing poverty. Thus for a given level of poverty, the better the policies and institutions, the larger the aid program should be. Fortunately, about 75 percent of the world's poor live in countries with policy environments good enough for aid to be effective in reducing poverty.

The third factor is the diminishing returns to aid. Even in environments with good policies and severe poverty, the amount of aid that can be effectively absorbed is limited—although the limit is high, about 20 percent of GDP.¹ But even in environments with poor policies, the first few million dollars of aid are worthwhile. Diminishing returns simply set in much sooner if policies are poor.

If donors had followed these three simple allocation criteria of poverty, policy, and diminishing returns, they could have helped considerably more people emerge from poverty (Collier and Dollar forthcoming).

Poverty-efficient aid allocation involves a straightforward relationship between policy and aid: for a given level of poverty, the amount of aid should be greater when policy is better. In other words, donors should condition aid on the level of policy. In contrast, aid for reform conditioned aid on *change* in policy. Because there is clearly more scope for improvement in policy when policy is worse, aid for reform would tend to bias aid flows toward weaker policy environments. This is indeed what happened in the 1990s (figure 1).

During the past decade aid tended to flow into policy environments that were too weak to use it effectively to reduce poverty—and tended to taper off in environments with policy good enough to use it effectively. This misallocation had serious

repercussions. Since there was only so much aid to go around, the people who paid the price for the large flows of aid to environments in which it was ineffective were those living in the better policy environments who could otherwise have been lifted out of poverty.

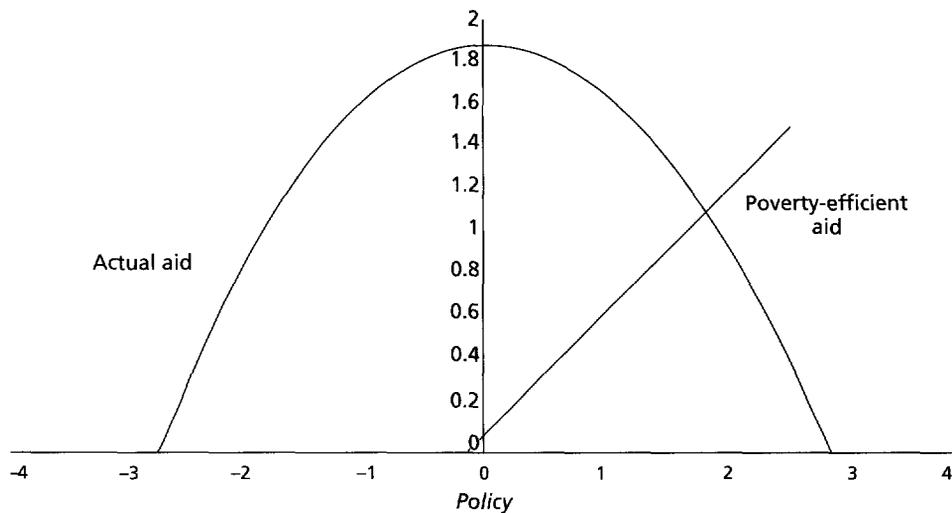
Conditioning aid on policy *change* rather than policy *level* leads not only to inefficient allocations between countries, but also to inefficient allocations over time. A democratic government that periodically faces an election typically will be more reluctant to change policies in the year before an election than in the year after an election. The cycle of elections thus creates a cycle of policy change. Aid flows conditioned on the rate of policy change become macroeconomically destabilizing: in the run-up to an election, just when a government is increasing its expenditures, its aid finance tends to be squeezed.

A Substitute for Government Commitment

Aid for reform may have weakened governments' capacity to work out and communicate their own strategies. Often governments had little real involvement in preparing the aid agreements they signed. The documents were sometimes prepared in Washington even before donor missions arrived in the country. Because governments knew that donor negotiators would try to coerce them into agreeing to do

Figure 1. Relationship between Policy and Poverty-Efficient and Actual Aid Allocation in the 1990s

Aid (percentage of GDP)



Note: The figure measures policy using the World Bank Country Policy and Institutional Assessment Index (CPIA). This index measures a broad range of policy, including macroeconomic, structural, and public sector management. Most countries fall into the range of -2 to +2. The CPIA is measured on the range 1-6, here transformed into a series centered on zero.

Source: Collier and Dollar forthcoming.

more than they seemed to want to do, it was reasonable for them to seem to want to do even less than they really wanted to do. The incentives of aid for reform thus impeded rather than assisted even the reforms the governments believed in.

Governments also had little incentive to sell policies to their electorates. An entire doctrine evolved in Washington about the efficacy of international financial institutions as scapegoats: governments could blame these institutions for unpopular but necessary policy changes. The Zimbabwean minister of information commented to the media on his government's economic reform program in 1997, "It's the IMF's program; we had to go along with it" (IMF 1998, p. 35). There was some truth to the idea that it can be helpful for governments to have a scapegoat. The Zimbabwean government might not have initially reformed so quickly if it had had to carry the electorate with it.

As a general proposition, however, the scapegoat theory is surely wrong. The role of scapegoat has a corollary—the electorate is seriously misinformed about key aspects of policy. Repeated often enough by the government, the message "this policy is dreadful but we are being forced to do it by foreigners" produces suspicion, defeatism, and confusion. Perhaps if the Zimbabwean government had had to carry the electorate along, it would not have reversed the reforms quite so readily the following year.

An Incentive for Governments to Reverse Reforms

Another possible effect of the attempt to coerce reform is the psychological phenomenon known as reactance. According to clinical psychologists, if someone tries to force you to do something, unless that person has total power over you your natural reaction is to do the opposite. Only by doing the opposite can you reestablish your freedom of action. Governments may have found it more attractive to reverse reforms because such action reestablished this freedom. Electorates may have gone along with the reversal because they had been told so often that the reforms were not a national choice.

What Is Needed for Sustainable Improvement in Policies and Institutions

Aid for reform was well intentioned, but it was based on a misunderstanding of how policies and institutions can be changed sustainably. At the risk of oversimplification, I suggest that we now know within reason what constitutes a good macroeconomic policy environment but have less idea about what constitutes good institutions.

Good macroeconomic policies are fairly generic. An overvalued exchange rate has qualitatively the same damaging effects in India as in Peru, and this is why such policy variables are significant in growth regressions (Burnside and Dollar 2000). Some high-profile issues are still in dispute, of course. For example, most scholars now probably accept that unrestricted capital accounts may be unwise on balance in many developing countries. But even this doctrine has nuances: despite controls, there has been so much capital flight in Africa that open capital accounts may need to be part

of the strategy to attract capital back. Good institutions are much more historically specific. We know, for example, that the design of legal institutions is much less important than their operation (Berkowitz, Pistor, and Richard 2000). Reforming institutions may require a different process than reforming macroeconomic policy.

Policy Reform

Policies depend largely on the balance between domestic political constituencies. Countries with very poor policies have large latent constituencies for policy change because poor policies inflict poor outcomes. Policy reform depends on strengthening the constituencies that suffer from the poor policies. The 1990s brought greater empowerment of these constituencies. As one example, the visible failure of the Soviet model stimulated a wave of democratization and provided a wealth of information for the debate on development policy.

Governments were pressured into improving their country's economic environment, though fitfully. Even benevolent leaders learned from having to listen more closely to their population. For example, when President Museveni of Uganda went out to rural areas to campaign for votes in the country's first fair presidential election, he discovered that what people wanted was free primary education for their children. This demand became so pressing that Museveni announced a massive change in policy in the middle of his campaign. The abolition of primary school fees led to a doubling in school enrollments the following year. We now know that pressure from civil society is effective in improving government performance. For example, civil liberties such as freedom of the press raise the return on public investment (Isham, Kaufmann, and Pritchett 1995).

It was easy for the Ugandan electorate to understand that they needed a policy change in primary education. But in many areas of policy electorates have too little information to discipline leaders effectively. Ironically, the slogan "It's the economy, stupid!" has described political debate in the poorest countries less accurately than in the richest. There are four major knowledge bottlenecks: lack of information, lack of a way to turn information into knowledge, lack of capacity for analysis, and lack of capacity for policy design.

To be effective, democracy needs the disclosure of *information*. Recently there has been a great deal of emphasis on the need for transparency in the banking system so that depositors can assess solvency. But the lack of transparency in government is a much larger problem. Electorates often lack basic information with which to assess the performance of their government.

Electorates often also lack a basis for comparison. Especially when neighboring governments perform just as badly as their own, electorates have no way to judge whether their government's performance could be better. This is why regional role models have been so valuable. The "gang of four" economies in East Asia, and Chile in Latin America, fostered a transformation across their regions. Africa and the Middle East have had no such models. Electorates need a window onto the world to turn information into knowledge.

The effects of many policies can be understood only through analysis. For example, an economist usually finds it fairly straightforward to work out the true effects of trade restrictions. But these effects are not obvious to electorates unless they have a remarkably sophisticated understanding of the economy or, more realistically, trusted authorities to explain the effects. Many developing countries lack such authorities. The consequence: supported by economic myths, poor policies persist. In a trivial but revealing example, the largest denomination of currency in Nigeria is barely worth a dollar, because people imagine (wrongly) that introducing higher-denomination notes would be inflationary. Moreover, the lack of high-denomination notes increases the costs of transactions because firms must buy note counting machines. It also increases the government's costs of printing money—the Nigerian government pays the equivalent of a real interest rate of more than 20 percent a year on its currency supply. Seigniorage—government revenue from supplying currency—far from being the government's cheapest form of debt, is its most expensive (Teriba 1998). The country needs a Nigerian think tank to puncture the myth that high-denomination notes are inflationary and a Nigerian press campaign to spread the message. More generally, the society needs to develop a *capacity for analysis*.

Many governments themselves lack the capacity to work out and communicate a coherent program of reform. As discussed above, under the aid-for-reform strategy governments may even have learned techniques of passive resistance to reform. Governments need a *capacity for design* of reform programs.

Institutional Reform

Now consider the process of institutional reform. Even simple, specific issues of appropriate institutional design are unresolved. For example, economists commonly lump together Anglo-Saxon economic institutions and contrast them with continental European or Japanese models. But even within the Anglo-Saxon model they disagree radically about which bankruptcy procedure is better—the U.S. model that uses the courts, which seems to be gaining popularity, or the British model that bypasses the courts and uses the private sector, which may be preferable where the courts are weak. The best policy for regulating utilities is even less clear. OECD countries keep changing between price caps and profit rate caps, neither of which is better in all situations. Institutions seem to matter, but good institutional blueprints are more difficult to identify than good macroeconomic policy blueprints.

Because knowledge in such circumstances becomes a constraint on reform, creating an effective reform process is synonymous with creating an effective knowledge discovery process. The core is experiment and competition. We need enough variety and choice of institutions to find out which are better than others. Companies could be allowed to choose a legal system when specifying a contract. Local governments could be encouraged to adopt innovative institutional arrangements. We know, for example, that U.S. states with elected boards overseeing their electricity utilities have persistently lower electricity prices than those with appointed boards. But we do not know

whether an elected judiciary would improve the functioning of courts in Africa, and we will never know until some African courts experiment with such an institution.

Implications for the Role of Donors

The donor community can do a great deal to assist both policy reform and institutional reform. To encourage policy reform, donors can help pro-reform constituencies within the bounds of appropriate conduct. The international community can legitimately encourage standards of good practice for information disclosure. Donors are obvious conduits for information on how performance differs elsewhere, and can supply the analysis that shows the true effects of policy. To encourage institutional reform, donors can show governments the range of potentially viable options available and stress that diversity and experiment are legitimate responses to uncertainty. They can finance pilot institutional reforms that, if successful, can be scaled up through imitation.

But the international community cannot supplant governments in the reform process. Governments cannot abdicate responsibility for working out a development strategy. Still, there are good reasons why this process should be a partnership between governments and the international community rather than an exclusively government activity. Governments will usually start from limited information, limited knowledge, and limited analytic and design capacity—the ostensible reason for the international financial institutions' heavy involvement in drawing up reform strategies. What is the difference, then, between collaboration based on partnership and collaboration based on coercion?

Cooperation—for Greater Information Sharing

How governments and donors interact determines what can be achieved. As noted, negotiation and collaboration do not coexist well. In a negotiation the government has an incentive to conceal information and convey misleading signals, while the donor team has an incentive to extract government concessions. In contrast, in a collaboration where the prime purpose of government-donor interaction is to build a common strategy for development, both can reap gains from pooling information. Because governments, international financial institutions, and donors have such different information advantages, such gains can be large. Cooperation should produce better-informed programs.

Consensus Building—for Faster Reform

Except in the short term, the pace of reform is likely to be constrained by the electorate's willingness to accept change. In Africa there is evidence that policy reversals are a response to urban rioting following changes that disadvantage urban population groups (Morrisson, Lafay, and Dessus 1994). Such policy instability is to no one's advantage. Of course, no government can govern entirely by consensus. But industrial countries achieve much of their policy change through persuasion, accommodation, and co-option. What is needed is a mechanism for reaching agreement ex

ante. If every reform is placed in the larger context of a medium-term strategy, each social group can more readily accept that it will lose from some policy changes but gain overall. It then becomes rational to sustain the integrity of the strategy rather than to block each change from which the group loses. Consensus building should produce faster sustainable policy change.

Governments will not always be able to build such consensus, which requires skill and a belief in the possibility of mutual gains. But governments that disown their economic reform programs as externally imposed build a consensus against reform. For institutional reform, where the absence of international consensus makes domestic diversity of views desirable, it may still be possible to build consensus around the need to experiment and to try competing approaches.

Government-Donor Cooperation—for More Stable Aid Flows

Donors' participation in designing strategy will increase the flow of information to them and thus reassure them of government intentions and the viability of government plans. This reassurance is important in determining not only the scale but also the stability of public resource flows. Donor financial support should be a source of stability: the government, international agencies, and investors should all be able to perceive such support as reliable over the medium term.

Aid for reform cumulatively undermined this perception of reliability. When governments breached their agreements on conditions, the continued flow of aid depended on donors' discretionary decisions to grant waivers. Aid flows were volatile and unpredictable as a result. One reason aid tends to taper off in good policy environments is that it is seen as so unreliable that people feel it is safer to learn to live with less of it. Associated with this is loose talk of the problem of aid dependence. As mentioned, the tapering off of aid denies it to precisely the environments in which it can be so highly effective in reducing poverty.

In fact, donors have not been unreliable over the past quarter of a century. In Africa aid has been less volatile than government revenue, so a large aid inflow has been a source of stability rather than instability (Collier 1999). But donor involvement placed in a cooperatively designed medium-term framework would be a firmer basis for donor commitment. In poor countries with reasonable policies and institutions, increasing aid flows will be desirable for poverty reduction for at least a decade. *Government-donor cooperation can produce larger and less volatile aid flows in environments where aid is most effective.*

Medium-Term Development Strategies—for Better Donor Coordination

Some governments are suspicious of coordination among donors because they see donors as ganging up on them to force through their own priorities. Agreement on objectives is therefore a necessary condition for coordination. The current lack of coordination reduces aid effectiveness. At the project level it results in duplication

of some interventions, omission of others, and occasional incompatible interventions. At the macro level it results in misallocation of aid among countries.

If all donors followed the poverty-efficient allocation rule of the International Development Association (IDA)—targeting aid based on a country's poverty and policies—they could substantially increase the number of people lifted out of poverty. But many donors do not (see figure 1). The Bank thus faces a dilemma. At one extreme it can stick to the rule for its IDA resources and encourage others to adopt it. At the other extreme it can use the rule as a guide for total aid resources, using IDA resources to smooth out the omissions of the rest of the donor community. Each of these extremes involves major difficulties. Medium-term development strategies can help to reduce the dilemma—by making aid provision more dependent on the total needs implied by a viable development path and less dependent on those implied by the provision of short-term incentives for policy change. *An agreed medium-term development strategy can improve donor coordination.*

Visible Political Consensus—for Reassuring Investors

Finally, private investors need reassurance. Africa is rated as the riskiest investment region in the world—even countries that have been strong reformers are rated as severely risky (Collier and Pattillo 2000). Aid-for-reform commitments have evidently failed to reassure private investors: the commitments lack credibility. A visible process of creating a medium-term social consensus around policy reform, if successful, would build investor confidence.

The Ugandan presidential election was the first substantial opportunity for Ugandan society to discuss a vision of the future, and it resulted in a large vote in favor of modernization. The election was followed by the largest improvement in investor risk ratings experienced by any African economy. In East Asia the opposite phenomenon occurred: the collapse in the risk ratings for Indonesia partly reflected the lack of political consensus. *Visible political consensus reassures investors more effectively than conditionality.*

Conclusion

Poverty reduction is now possible on a grand scale if donors and governments can navigate the policy and institutional changes needed for broad-based growth. In many social contexts this is likely to be easier if there is an informed constituency and if the government attempts to build consensus around the reforms. Because providing information and building consensus sound as wholesome as motherhood and apple pie, it is easy to dismiss them as decorative rather than functional parts of the development business. I have tried to show why this would be a mistake. In the past we may have paid lip service to information provision and consensus building within society, but we have neglected them in practice. Aid for reform tried to bypass consensus building and led to governments publicly disowning their own

programs. We should not be surprised that private investors regarded the reform process as lacking credibility under these circumstances.

Developed societies do not just happen to have good policies and institutions. They have them because governments are pinioned to them by informed and engaged social groups. Somehow the international community needs to encourage the formation of this equilibrium in developing countries. I have described steps the international community could take to do this—working with governments to identify a critical path of policy reform and initiating institutional experiments and competitions to discover which institutions work best. I have also described why ex ante social consensus would both speed these processes and enhance their credibility. What I have described is my understanding of the Comprehensive Development Framework.

Notes

1. Collier and Dollar (forthcoming) estimate the rate of diminishing returns to aid for different policy environments. They measure aid at purchasing power parity exchange rates, which differ from actual exchange rates in aid-receiving countries on average by a factor of around three. The 20 percent limit used here reflects an adjustment back to actual exchange rates, by which aid is normally measured.

References

- Appleton, Simon. 2001. "Education in Uganda: What Can We Expect from Universal Primary Education?" In Ritva Reinikka and Paul Collier, eds., *Uganda's Recovery: The Role of Farms, Firms, and Government*. Washington, D.C.: World Bank.
- Atingi-Ego, M., and Louis Kasekende. 1997. "Financial Liberalization and Its Implications for the Domestic Financial System: The Case of Uganda." African Economic Research Consortium, Nairobi.
- Berkowitz, Daniel, Katharina Pistor, and Jean-François Richard. 2000. "Economic Development, Legality, and the Transplant Effect." CID Working Paper 1. Harvard University, Center for International Development, Cambridge, Mass.
- Bikhchandani, Sushil, David Hirschleifer, and Ivo Welch. 1998. "Learning from the Behavior of Others: Conformity, Fads and Informational Cascades." *Journal of Economic Perspectives* 12 (3): 151–70.
- Burnside, Craig, and David Dollar. 2000. "Aid, Policies and Growth." *American Economic Review* 90(4): 847–68.
- Collier, Paul. 1997. "The Failure of Conditionality." In Catherine Gwin and J. M. Nelson, eds., *Perspectives on Development*. Washington, D.C.: Overseas Development Council.
- . 1999. "Aid Dependency: A Critique." *Journal of African Economies* 8 (4): 528–45.
- Collier, Paul, and Ashish Garg. 1999. "Kin Groups and Wages in the Ghanaian Labour Market." *Oxford Bulletin of Economics and Statistics* 61 (2): 133–51.
- Collier, Paul, and Catherine Pattillo, eds. 2000. *Investment and Risk in Africa*. London: Macmillan.
- Collier, Paul, and David Dollar. Forthcoming. "Aid Allocation and Poverty Reduction." *Europa Economic Review*.
- Collier, Paul, Anke Hoeffler, and Catherine Pattillo. 2001. "Flight Capital as a Portfolio Choice." *World Bank Economic Review* 15 (1): 55–80.

- Cox, Donald, and Emmanuel Jimenez. 1995. "Private Transfers and the Effectiveness of Public Income Redistribution in the Philippines." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence*. Baltimore, Md.: Johns Hopkins University Press.
- Devarajan, Shantayanan, David Dollar, and Torgny Holmgren, eds. 2001. *Aid and Reform in Africa*. Washington, D.C.: World Bank.
- Dollar, David, and Aart Kraay. 2000. "Growth Is Good for the Poor." World Bank, Development Research Group, Washington, D.C.
- Dollar, David, and Jakob Svensson. 2000. "What Explains the Success or Failure of Structural Adjustment Programmes?" *Economic Journal* 110(466): 894–917.
- Easterly, William and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112 (4): 1203–50.
- Feyzioglu, Tarhan, Vinaya Swaroop, and Min Zhu. 1997. "Foreign Aid's Impact on Public Spending." Policy Research Working Paper 1610. World Bank, Policy Research Group, Public Economics, Washington, D.C.
- IMF (International Monetary Fund). 1998. *External Evaluation of the ESAF*. Washington, D.C.
- Isham, Jonathan, and Daniel Kaufmann. 1999. "The Forgotten Rationale for Policy Reform: The Productivity of Investment Projects." *Quarterly Journal of Economics* 114: 149–84.
- Isham, Jonathan, Daniel Kaufmann, and Lant Pritchett. 1995. "Government and the Returns on Investment." Policy Research Working Paper 1550. World Bank, Policy Research Group, Poverty and Human Resources, Washington, D.C.
- Li, Hongyi, Lyn Squire, and Heng-fu Zou. 1998. "Explaining International and Intertemporal Variations in Income Inequality." *Economic Journal* 108 (446): 26–43.
- McGee, Rosemary. 2000. "Participatory Poverty Assessment and Household Survey Findings on Poverty Trends in Uganda." Institute for Development Studies, Brighton, U.K.
- Morrison, Christian, Jean-Dominique Lafay, and Sébastien Dessus. 1994. "Adjustment Programs and Politico-Economic Interactions in Developing Countries." In G. A. Cornia and G. K. Helleiner, eds., *From Adjustment to Development in Africa: Conflict, Controversy, Convergence, Consensus?* London: Palgrave.
- Reinikka, Ritva, and Jakob Svensson. 1999. "How Inadequate Provision of Public Infrastructure and Services Affects Private Investment." Policy Research Working Paper 2262. World Bank, Development Research Group, Public Economics, and Macroeconomics and Growth, Washington, D.C.
- Svensson, Jakob. 1999. "Why Conditional Aid Doesn't Work and What Can Be Done About It." World Bank, Development Research Group, Washington, D.C.
- Teriba, Ayodeke O. 1998. "Business and Policy Issues in Re-Denominating the Naira." *LBS Management Review* 3 (2): 101–25.
- Thaker, Strom. 1999. "The High Politics of IMF Lending." *World Politics* 52: 38–75.
- Uganda, Ministry of Finance and Planning. 1999. "Participatory Poverty Assessment." Kampala.

Development Strategies for the 21st Century

Dani Rodrik

The lesson of the 20th century is that successful development requires markets underpinned by solid public institutions—institutions that protect property rights, regulate market participants, maintain macroeconomic stability, provide social insurance, and manage conflict. A variety of institutional setups could serve these functions, but any imported blueprints should be filtered through local practice and needs. International rules and the loan conditionality imposed by international financial institutions ought to leave room for development policies to diverge from the dominant orthodoxies. Today's advanced industrial countries owe their success to having developed their own workable models of a mixed economy. Developing nations need to fashion their own brands. Economic development will ultimately derive from homegrown strategies, not imitation of U.S.-style capitalism.

The idea of a mixed economy is possibly the most valuable heritage that the 20th century bequeathed to the 21st in the realm of economic policy. The 19th century discovered capitalism. The 20th learned how to tame it and render it more productive by supplying the institutional ingredients of a self-sustaining market economy: central banking, stabilizing fiscal policy, antitrust and regulation, social insurance, political democracy. It was during the 20th century that these elements of the mixed economy took root in today's advanced industrial countries. The simple idea that markets and the state are complements—recognized in practice if not always in principle—made possible the unprecedented prosperity that the United States, Western Europe, and parts of East Asia experienced during the second half of the century.

Dani Rodrik is professor of international political economy at Harvard University's John F. Kennedy School of Government. The first draft of this paper was presented at the conference *Developing Economies in the 21st Century: The Challenges to Globalization*, organized by the Institute of Developing Economies–Japan External Trade Organization, on 26–27 January 2000 in Chiba, Japan. The author is grateful to IDE for financial support provided under a research agreement on *Reconsideration of Development Strategies toward the 21st Century*, and to Nancy Birdsall, François Bourguignon, and Fukunari Kimura for helpful comments.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

The truism that private initiative and collective action are both required for economic success arrived in developing countries rather late. As most of them were becoming independent in the 1950s and 1960s, the apparently successful example of the Soviet Union and the antimarket ideology of national governing elites resulted in heavily state-centric development strategies. In Latin America, where countries had long been independent, the dominant “structuralist” view held that market incentives would fail to elicit much of a supply response. Throughout the developing world the private sector was regarded with skepticism, and private initiative was severely circumscribed.

These views underwent a radical transformation during the 1980s under the joint influence of a protracted debt crisis and the teachings of the Bretton Woods institutions. The “Washington consensus,” emphasizing privatization, deregulation, and trade liberalization, was embraced enthusiastically by policymakers in Latin America and post-socialist Eastern Europe. The reception was more guarded and cautious in Africa and Asia, but there too policies took a decided swing toward markets. The market-oriented reforms at first paid little attention to institutions and the complementarity between the private and public spheres of the economy. The role assigned to the government did not go beyond that of maintaining macroeconomic stability and providing education. The priority was on rolling back the state, not on making it more effective.

A more balanced view began to emerge during the closing years of the 20th century as it became clearer that the Washington consensus would fail to deliver on its promise. The talk in Washington turned toward second-generation reforms, governance, and “reinvigorating the state’s capability” (World Bank 1997, p. 27). And multilateral institutions began to take a considerably humbler view of conditionality. Several developments added fuel to the discontent over the orthodoxy. The first of these was the dismal failure in the Russian Federation of price reform and privatization in the absence of a supportive legal, regulatory, and political apparatus. The second was the widespread dissatisfaction with market-oriented reforms in Latin America and the growing realization that these reforms had paid too little attention to mechanisms of social insurance and to safety nets. The third and most recent was the East Asian financial crisis, which exposed the dangers of allowing financial liberalization to run ahead of adequate regulation.

So we enter the 21st century with a better understanding of the complementarity between markets and the state—a greater appreciation of the virtues of the mixed economy. That is the good news. The bad news is that the operational implications of this understanding for the design of development strategies are not that clear. There remains plenty of opportunity for mischief on the policy front. In particular, it is unlikely that an augmented Washington consensus strategy—appending to the old orthodoxy a new set of blueprints aimed at so-called second-generation reforms—will take us very far. As I argue below, the state and the market can be combined in diverse ways. There are many different models of a mixed economy. The major challenge facing developing nations in the first decades of the 21st century is to fashion their own brands of the mixed economy.

In what follows, I review some of the principles that should guide this quest. I begin by providing a capsule history of the post–World War II growth performance of developing countries. Since the reasons for the disappointing growth performance since the late 1970s are intricately linked with current policy prescriptions, I present my own interpretation of what went wrong. This interpretation underscores the importance of domestic institutions and downplays the role of microeconomic factors (including trade policy) in the post-1980 growth collapse.

In a more detailed analysis of market-supporting institutions, I discuss five functions that public institutions must serve for markets to work adequately: protection of property rights, market regulation, macroeconomic stabilization, social insurance, and conflict management. In principle, a large variety of institutional setups could serve these functions. We need to be skeptical of the notion that a specific institution observed in a country (the United States, say) is the type that is most compatible with a well-functioning market economy.

Partial and gradual reforms have often worked better because reform programs that are sensitive to institutional preconditions are more likely to be successful than those that assume that new institutions can be erected wholesale overnight. Learning and imitation from abroad are important elements of a successful development strategy. But imported blueprints need to be filtered through local experience and deliberation.

What are some of the implications of this for international governance? A key conclusion is that international rules and the conditionality imposed by international financial institutions ought to leave room for development policies that diverge from the dominant orthodoxies of the day. Trade and capital flows are important insofar as they allow developing countries access to cheaper capital goods. But the links between opening to trade and capital flows and subsequent growth are weak, uncertain, and mediated through domestic institutions.

Some Lessons of Recent Economic History

Many developing countries experienced unprecedented rates of economic growth during the postwar period until the late 1970s.¹ More than 40 of these countries grew at annual per capita rates exceeding 2.5 percent until the first oil shock hit. At this rate of growth incomes double every 28 years or less—that is, every generation. The list of countries with this enviable record goes far beyond the handful of usual East Asian suspects and covers all parts of the globe: 12 countries in Latin America, 6 in the Middle East and North Africa, and even 15 in Sub-Saharan Africa (Rodrik 1999a, table 4.1). There can be little doubt that economic growth led to substantial improvements in the living conditions of the vast majority of households in these countries.

The Role of Import Substitution Policies

Most of the countries that did well in this period followed import substitution policies. These policies spurred growth and created protected and therefore profitable

home markets for domestic entrepreneurs to invest in. Contrary to conventional wisdom, growth driven by import substitution did not produce tremendous inefficiencies on an economywide scale. In fact, the productivity performance of many Latin American and Middle Eastern countries was, in a comparative perspective, exemplary (Rodrik 1999a, table 4.2). In 1960–73 several countries in Latin America (for example, Brazil, the Dominican Republic, and Ecuador), the Middle East and North Africa (the Islamic Republic of Iran, Morocco, and Tunisia), and Sub-Saharan Africa (Côte d’Ivoire and Kenya) experienced more rapid growth in total factor productivity than any of the East Asian economies (with the possible exception of Hong Kong, for which comparable data are not available). Mexico, Bolivia, Panama, Egypt, Algeria, Tanzania, and Zaire experienced faster productivity growth than all but Taiwan, China. Productivity growth estimates of this type are not without serious problems, and one can quibble with the methodologies employed. Nevertheless, there is little reason to believe that the estimates of Collins and Bosworth (1996), from which these numbers are drawn, are seriously biased in the way that they rank different countries.

Thus as a strategy of industrialization, intended to raise domestic investment and enhance productivity, import substitution apparently worked pretty well in a broad range of countries until at least the mid-1970s. Despite its problems, import substitution achieved a more than respectable record. Had the world come to an end in 1973, the policy would not have acquired its dismal reputation, nor would East Asia have deserved its “miracle” appellation.

Collapse of Growth

Following the oil shock of 1973, however, things began to look very different. The median growth rate for developing countries fell from 2.6 percent in 1960–73 to 0.9 percent in 1973–84 and to 0.8 percent in 1984–94. The dispersion in performance across developing countries increased sharply, with the coefficient of variation for national growth rates increasing threefold after 1973 (Rodrik 1999a, table 4.3). The Middle East and Latin America, which had led the developing world in total factor productivity growth before 1973, not only fell behind but experienced negative productivity growth on average thereafter. In Sub-Saharan Africa, where productivity growth had been undistinguished but still positive, it also turned negative. Only East Asia held its own in productivity growth, while South Asia actually improved its performance.

Were these economic downturns a result of the “exhaustion” of import substitution policies, whatever that term means? The common timing implicates instead the turbulence experienced in the world economy following 1973—the abandonment of the Bretton Woods system of fixed exchange rates, two major oil shocks, various other commodity boom-and-bust cycles, and the Volcker interest rate shock of the early 1980s. The fact that some of the most ardent followers of import substitution policies in South Asia—India and Pakistan in particular—managed to either hold on to their growth rates after 1973 (Pakistan) or increase them (India) also suggests that more than just import substitution was involved.

The actual story is straightforward. The proximate reason for the economic collapse was the inability to adjust macroeconomic policies appropriately in the wake of these external shocks. Macroeconomic maladjustment gave rise to a range of syndromes associated with macroeconomic instability—high or repressed inflation, foreign exchange scarcity and large black market premia, external payments imbalances, and debt crises—that greatly magnified the real costs of the shocks. Indeed, there was a strong association between inflation and black market premia and the magnitude of economic collapse experienced in different countries. The countries that suffered the most were those with the largest increases in inflation and black market premia for foreign currency (Rodrik 1999a, figure 4.1). The culprits were poor monetary and fiscal policies and inadequate adjustments in exchange rate policy, sometimes aggravated by shortsighted policies of creditors and the Bretton Woods institutions. Trade and industrial policies had very little to do with bringing on the crisis.

Why were some countries quicker to adjust their macroeconomic policies than others? The deeper determinants of growth performance after the 1970s are rooted in the ability of domestic institutions to manage the distributional conflicts triggered by the external shocks of the period.

Think of an economy that is suddenly and unexpectedly confronted with a drop in the price of its main export (or a sudden reversal of capital flows). The textbook prescription for this economy is a combination of expenditure-switching and expenditure-reducing policies—that is, a devaluation and fiscal retrenchment. But the precise manner in which these policy changes are administered can have significant distributional implications. Should the devaluation be accompanied by wage controls? Should import tariffs be raised? Should the fiscal retrenchment take place through spending cuts or tax increases? If spending is to be cut, what types of expenditures should bear the brunt of the cuts? Should interest rates be raised to rein in private spending as well?

In general, macroeconomic theory does not have a clear preference among the available options. But since each of the options has predictable distributional consequences, in practice much depends on the severity of the social conflicts that lie beneath the surface. If the appropriate adjustments can be undertaken without causing an outbreak of distributional conflict or upsetting prevailing social bargains, the shock can be managed with few long-lasting effects on the economy. If they cannot be, the economy could be paralyzed for years as inadequate adjustment condemns the country to foreign exchange bottlenecks, import compression, debt crises, and bouts of high inflation. Furthermore, deep social divisions provide an incentive to governments to delay needed adjustments and to take on excessive levels of foreign debt, in the expectation that other social groups can be made to pay the eventual costs.

In short, social conflicts and their management play a key role in transmitting the effects of external shocks to economic performance. Societies with deep social cleavages and poor institutions of conflict management tend not to be very good at handling shocks. In such societies the economic costs of exogenous shocks—such as deterioration in the terms of trade—are magnified by the distri-

butional conflicts that are triggered. Such conflicts diminish the productivity with which a society's resources are utilized—by delaying needed adjustments in fiscal policies and key relative prices (such as the real exchange rate or real wages) and by diverting activities away from the productive and entrepreneurial spheres. Cross-national evidence is supportive of this argument: macroeconomic disequilibrium and growth collapse have been more likely in countries with high degrees of income inequality and ethnolinguistic fragmentation, and less likely in countries with democratic institutions or high-quality public institutions (Rodrik 1999b).

Lessons in Conflict Management from the East Asian Financial Crisis

The same logic played out in the recent East Asian financial crisis. One lesson of the crisis is that international capital markets do a poor job of discriminating between good and bad risks. It is hard to believe that there was much collective rationality in investor behavior before and during the crisis: financial markets got it badly wrong either in 1996, when they poured money into the region, or in 1997, when they pulled back en masse. The implication is that relying excessively on liquid, short-term capital (as all three of the worst affected countries did) is a dangerous strategy.

A second lesson is that trade orientation in itself has little to do with the propensity to be hit with severe liquidity problems. The Asian economies most affected by the reversal in capital flows were among the most outward-oriented economies in the world, routinely pointed out as examples for other countries to follow. The determinants of the crisis—as with the debt crisis of 1982 and the Mexican peso crisis of 1994—were financial and macroeconomic. Trade and industrial policies were, at best, secondary.²

A third lesson of the East Asian crisis is that domestic institutions of conflict management are critical in containing the adverse economic consequences of the initial shock. Indonesia, an ethnically divided society ruled by an autocracy, eventually descended into chaos. The democratic institutions of the Republic of Korea and Thailand, with their practices of consultation and cooperation among social partners, proved much more adept at generating the requisite policy adjustments. This recent experience has demonstrated once again the importance of institutions—particularly democratic institutions—in dealing with external shocks.

While democratic institutions are relatively recent in Korea and Thailand, they helped these two countries adjust to the crisis in several ways. First, they facilitated a smooth transfer of power from a discredited set of politicians to a new group of government leaders. Second, democracy imposed mechanisms of participation, consultation, and bargaining, enabling policymakers to fashion the consensus needed to undertake the necessary policy adjustments decisively. Third, because democracy provides for institutionalized mechanisms of “voice,” the Korean and Thai institutions obviated the need for riots, protests, and other kinds of disruptive

actions by affected groups and reduced support for such behavior from other groups in society.

A Different Interpretation

Many of the lessons that the development community has internalized from recent economic history are in need of revision. In my view the correct interpretation goes something like this. First, import substitution worked rather well for about two decades. It led to higher investment rates and unprecedented economic growth in scores of countries in Latin America, in the Middle East and North Africa, and even some in Sub-Saharan Africa. Second, when the economies of these same countries began to fall apart in the second half of the 1970s, the reasons had little to do with import substitution policies or the extent of government interventionism. The countries that weathered the storm were those in which governments undertook the appropriate macroeconomic adjustments (in fiscal, monetary, and exchange rate policy) rapidly and decisively. Third, and more fundamentally, success in adopting these macroeconomic adjustments was linked to deeper social determinants. It was the ability to manage the domestic social conflicts triggered by the turbulence of the world economy during the 1970s that made the difference between continued growth and economic collapse. Countries with deeper social divisions and weaker institutions (particularly of conflict management) experienced greater economic deterioration in response to the external shocks of the 1970s.

Taken together, these points provide an interpretation of recent economic history that differs from much current thinking. By emphasizing the importance of social conflicts and institutions—at the expense of trade strategy and industrial policies—they also suggest quite a different perspective on development policy. If I am right, the main difference between Latin American countries, say, and East Asian economies was not that those in the first group remained closed and isolated while those in the second integrated themselves with the world economy. The main difference was that Latin American countries did a much worse job of dealing with the turbulence emanating from the world economy. The countries that got into trouble were those that could not manage openness, not those that were insufficiently open.

A Taxonomy of Market-Sustaining Public Institutions

Institutions do not figure prominently in the training of economists.³ The standard Arrow-Debreu model, with its full set of complete and contingent markets extending indefinitely into the future, seems to require no assistance from nonmarket institutions. *But of course this is quite misleading even in the context of that model.* The standard model assumes a well-defined set of property rights. It also assumes that contracts are signed with no fear that they will be revoked when it suits one of the parties. So in the background there exist institutions that establish and protect property rights and enforce contracts. There has to be a system of laws and courts to make even “perfect” markets function.

Laws, in turn, have to be written, and they have to be backed by the use of sanctioned force. That implies a legislator and a police force. The legislator's authority may derive from religion, family lineage, or access to superior violence, but in each case she needs to ensure that she provides her subjects with the right mix of "ideology" (a belief system) and threat of violence to forestall rebellion from below. Or the legislator's authority may derive from the legitimacy provided by popular support, in which case she needs to be responsive to her constituency's (voters') needs. In either case we have the beginnings of a governmental structure that goes well beyond the narrow needs of the market.

One implication of all this is that the market economy is necessarily "embedded" in a set of nonmarket institutions. Another is that not all these institutions are there to serve the needs of the market economy first and foremost, even if their presence is required by the internal logic of private property and contract enforcement. The fact that a governance structure is needed to ensure that markets can do their work does not imply that the governance structure serves only that end. Nonmarket institutions will sometimes produce outcomes that are socially undesirable, such as the use of public office for private gain. They may also produce outcomes that restrict the free play of market forces in pursuit of larger goals, such as social stability and cohesiveness.

The rest of this section discusses five types of market-supporting institutions: property rights, regulatory institutions, institutions for macroeconomic stabilization, institutions for social insurance, and institutions of conflict management.

Property Rights

It is possible to envisage a thriving socialist market economy in theory, as the famous debates of the 1920s established. But today's prosperous economies have all been built on the basis of private property. As North and Thomas (1973) and North and Weingast (1989), among many others, have argued, the establishment of secure and stable property rights was a key element in the rise of the West and the onset of modern economic growth. Entrepreneurs do not have the incentive to accumulate and innovate unless they have adequate control over the return to the assets that are thereby produced or improved.

Note that the key word is *control* rather than *ownership*. Formal property rights do not count for much if they do not confer control rights. By the same token, sufficiently strong control rights may work adequately even in the absence of formal property rights. Russia today represents a case where shareholders have property rights but often lack effective control over enterprises. Town and village enterprises in China are an example where control rights have spurred entrepreneurial activity despite the absence of clearly defined property rights. As these examples illustrate, establishing "property rights" is rarely a matter of just passing a piece of legislation. Legislation in itself is neither necessary nor sufficient for the provision of secure control rights. In practice, control rights are upheld by a combination of legislation, private enforcement, and custom and tradition. They may be distributed more nar-

rowly or more diffusely than property rights. Stakeholders can matter as much as shareholders.

Moreover, property rights are rarely absolute, even when set formally in the law. My right to keep my neighbor out of my orchard does not normally extend to a right to shoot him if he enters it. Other laws or norms—such as those against murder—may trump property rights. Each society decides for itself the scope of allowable property rights and the acceptable restrictions on their exercise. Intellectual property rights are protected assiduously in the United States and most other advanced societies, but are not in many developing countries. But zoning and environmental legislation restricts the ability of households and enterprises in the rich countries to do as they please with their “property” to a much greater extent than is the case in developing countries. All societies recognize that private property rights can be curbed if doing so serves a greater public purpose. It is the definition of what constitutes “greater public purpose” that varies.

Regulatory Institutions

Markets fail when participants engage in fraudulent or anticompetitive behavior. They fail when transaction costs prevent the internalizing of technological and other nonpecuniary externalities. And they fail when incomplete information results in moral hazard and adverse selection. Economists recognize these failures and have developed the analytical tools required to think systematically about their consequences and about possible remedies. Theories of the second best, imperfect competition, agency, mechanism design, and many others offer an almost embarrassing choice of regulatory instruments to counter market failures. Theories of political economy and public choice offer cautions against unqualified reliance on these instruments.

In practice, every successful market economy is overseen by a panoply of regulatory institutions, regulating conduct in markets for goods, services, labor, assets, and finance. A few acronyms from the United States suffice to give a sense of the range of institutions involved: EPA, FAA, FCC, FDIC, FTC, OSHA, SEC. In fact, the freer are the markets, the greater is the burden on the regulatory institutions. It is no coincidence that the United States has the world’s freest markets as well as its toughest antitrust enforcement. It is hard to envisage a hugely successful high-technology company like Microsoft being dragged through the courts for alleged anticompetitive practices in any country other than the United States.

The lesson that market freedom requires regulatory vigilance was driven home recently by the experience in East Asia. In Korea and Thailand, as in so many other developing countries, financial liberalization and capital account opening led to financial crisis precisely because of inadequate prudential regulation and supervision.⁴

Regulatory institutions may need to extend beyond the standard list covering antitrust, financial supervision, securities regulation, and a few others. This is true especially in developing countries, where market failures may be more pervasive and the requisite market regulations more extensive. Recent models of coordination

failure and capital market imperfections make it clear that strategic government interventions may often be required to get out of low-level traps and elicit desirable private investment responses (see Hoff and Stiglitz 2001 for a useful survey and discussion). The experience of Korea and Taiwan, China, in the 1960s and 1970s can be interpreted in that light. The extensive subsidization and government-led coordination of private investment in these two economies played a crucial role in setting the stage for self-sustaining growth (Rodrik 1995). Many other countries have tried and failed to replicate these institutional arrangements.

And even Korea may have taken a good thing too far by maintaining the cozy institutional links between the government and *chaebol* well into the 1990s, at which point these links may have become dysfunctional. Once again, the lesson is that desirable institutional arrangements vary and that they vary not only across countries but also within countries over time.

Institutions for Macroeconomic Stabilization

Since Keynes, we have come to a better understanding of the reality that capitalist economies are not necessarily self-stabilizing. Keynes and his followers worried about shortfalls in aggregate demand and the resulting unemployment. More recent views of macroeconomic instability stress the inherent instability of financial markets and its transmission to the real economy. All advanced economies have come to acquire fiscal and monetary institutions that perform stabilizing functions, having learned the hard way about the consequences of not having them. Probably most important among these institutions is a lender of last resort—typically the central bank—which guards against self-fulfilling banking crises.

There is a strong current within macroeconomics thought, represented in its theoretically most sophisticated version by the real business cycles approach, that disputes the possibility or effectiveness of stabilizing the macroeconomy through monetary and fiscal policies. There is also a sense in policy circles, particularly in Latin America, that fiscal and monetary institutions—as currently configured—have added to macroeconomic instability, rather than reduced it, by following procyclical rather than anticyclical policies (Hausmann and Gavin 1996). These developments have spurred the trend toward central bank independence and helped open a new debate on designing more robust fiscal institutions.

Some countries—Argentina being the most significant example—have given up on a domestic lender of last resort altogether by replacing their central bank with a currency board. The Argentine calculation is that having a central bank that can occasionally stabilize the economy is not worth running the risk that the central bank will mostly *destabilize* it. Argentine history gives plenty of reason to think that this is not a bad bet. But can the same be said for Brazil or Mexico, or, for that matter, Indonesia or Turkey? A substantial real depreciation of the rupee, engineered through nominal devaluations, was a key ingredient of India's superlative economic performance during the 1990s. What might work for Argentina might not work for others. The debate over currency boards and dollarization illustrates the obvious,

but occasionally neglected, fact that the institutions needed by a country are not independent of that country's history.

Institutions for Social Insurance

A modern market economy is one in which change is constant and idiosyncratic (individual-specific) risk to incomes and employment is pervasive. Modern economic growth entails a transition from a static economy to a dynamic one in which the tasks that workers perform are in constant evolution and movement up and down the income scale is frequent. One of the liberating effects of a dynamic market economy is that it frees individuals from their traditional entanglements—the kinship group, religious organizations, the village hierarchy. The flip side is that it uproots them from traditional support systems and risk-sharing institutions. Gift exchanges, the fiesta, and kinship ties—to cite just a few of the social arrangements for equalizing the distribution of resources in traditional societies—lose much of their social insurance function. And as markets spread, the risks that have to be insured against become much less manageable in the traditional manner.

The huge expansion of publicly provided social insurance programs during the 20th century is one of the most remarkable features of the evolution of advanced market economies. In the United States it was the trauma of the Great Depression that paved the way for the major institutional innovations in this area: social security, unemployment compensation, public works, public ownership, deposit insurance, and legislation favoring unions (see Bordo, Goldin, and White 1998). As Jacoby (1998) notes, before the Great Depression the middle classes were generally able to self-insure or to buy insurance from private intermediaries. As these private forms of insurance collapsed, the middle classes threw their considerable political weight behind the extension of social insurance and the creation of what would later be called the welfare state. In Europe the roots of the welfare state reach in some cases back to the tail end of the 19th century. But the striking expansion of social insurance programs, particularly in the smaller economies most open to foreign trade, was a post-World War II phenomenon (Rodrik 1998b). Despite considerable political backlash against the welfare state since the 1980s, neither the United States nor Europe has significantly scaled back these programs.

Social insurance need not always take the form of transfer programs paid for out of fiscal resources. The East Asian model, represented well by the Japanese case, is one in which social insurance is provided through a combination of enterprise practices (such as lifetime employment and enterprise-provided social benefits), sheltered and regulated sectors (mom-and-pop stores), and an incremental approach to liberalization and external opening. Certain aspects of Japanese society that seem inefficient to outside observers—such as the preference for small-scale retail stores or extensive regulation of product markets—can be viewed as substitutes for the transfer programs that would otherwise have to be provided by a welfare state (as they are in most European nations).

An important implication of such complementarities among different institutional arrangements in a society is that it is very difficult to alter national systems in

a piecemeal fashion. One cannot (or should not) ask the Japanese to get rid of their lifetime employment practices or inefficient retail arrangements without ensuring that alternative safety nets are in place. Another implication is that substantial institutional changes come only in the aftermath of large dislocations, such as those created by the Great Depression or World War II.

Social insurance legitimizes a market economy because it renders it compatible with social stability and social cohesion. At the same time, the existing welfare states in Western Europe and the United States engender a number of economic and social costs—mounting fiscal outlays, an “entitlement” culture, long-term unemployment—which have become increasingly apparent. Partly because of that, developing countries, such as those in Latin America that adopted the market-oriented model following the debt crisis of the 1980s, have not paid sufficient attention to creating institutions of social insurance. The upshot has been economic insecurity and a backlash against the reforms. How these countries will maintain social cohesion in the face of large inequalities and volatile outcomes, both of which are being aggravated by the growing reliance on market forces, is a question without an obvious answer at the moment. But if Latin America and the other developing regions are to carve a different path in social insurance than that followed by Europe or North America, they will have to develop their own visions—and their own institutional innovations—to ease the tension between market forces and the yearning for economic security.

Institutions of Conflict Management

Societies differ in their cleavages. Some are made up of an ethnically and linguistically homogenous population marked by a relatively egalitarian distribution of resources (Finland?). Others are characterized by deep cleavages along ethnic or income lines (Nigeria?). These divisions hamper social cooperation and prevent the undertaking of mutually beneficial projects. Social conflict is harmful both because it diverts resources from economically productive activities and because it discourages such activities through the uncertainty it generates. Economists have used models of social conflict to shed light on such questions as the following: Why do governments delay stabilization when delay imposes costs on all groups (Alesina and Drazen 1991)? Why do countries rich in natural resources often do worse than countries that are resource-poor (Tornell and Lane 1999)? Why do external shocks often lead to protracted economic crises that are out of proportion to the direct costs of the shocks themselves (Rodrik 1999b)?

All these can be thought of as instances of coordination failure, in which social factions fail to coordinate on outcomes that would be of mutual benefit. Healthy societies have a range of institutions that make such colossal coordination failures less likely. The rule of law, a high-quality judiciary, representative political institutions, free elections, independent trade unions, social partnerships, institutionalized representation of minority groups, and social insurance are examples of such institutions. What makes these arrangements function as institutions of conflict man-

agement is that they entail a double “commitment technology”—they warn the potential “winners” of social conflict that their gains will be limited and assure the “losers” that they will not be expropriated. These arrangements tend to increase the incentives for social groups to cooperate by reducing the payoff to socially uncooperative strategies.

What Role for Institutional Diversity?

As the previous section shows, a market economy relies on a wide array of non-market institutions that perform regulatory, stabilizing, and legitimizing functions. Once these institutions are accepted as part and parcel of a market-based economy, traditional dichotomies between market and state or laissez-faire and interventionism begin to make less sense. These are not competing ways of organizing a society’s economic affairs; they are complementary elements that render the system sustainable. Every well-functioning market economy is a mix of state and market, laissez-faire and interventionism.

Another implication of the discussion in the previous section is that the institutional basis for a market economy is not uniquely determined. Formally, there is no single mapping between the market and the set of nonmarket institutions required to sustain it. This finds reflection in the wide variety of regulatory, stabilizing, and legitimizing institutions that we observe in today’s advanced industrial societies. The U.S. style of capitalism is very different from the Japanese style. Both differ from the European style. And even within Europe there are large differences between the institutional arrangements in, say, Germany and Sweden.

It is a common journalistic error to suppose that one set of institutional arrangements must dominate the others in overall performance. Thus the fads of the decade: With its low unemployment, high growth, and thriving culture, Europe was the continent to emulate throughout much of the 1970s. During the trade-conscious 1980s Japan became the exemplar of choice. And the 1990s were the decade of U.S.-style freewheeling capitalism. It is anybody’s guess which set of countries will capture the imagination if and when a substantial correction hits the U.S. stock market.⁵

The point about institutional diversity has a more fundamental implication. The institutional arrangements that we observe in operation today, varied as they are, constitute a subset of the full range of institutional possibilities. This is a point that has been forcefully and usefully argued by Unger (1998). There is no reason to suppose that modern societies have already managed to exhaust all the useful institutional variations that could underpin healthy and vibrant economies. Even if we accept that market-based economies require certain types of institutions, as listed in the previous section,

such imperatives do not select from a closed list of institutional possibilities. The possibilities do not come in the form of indivisible systems, standing or falling together. There are always alternative sets of arrangements capable of meeting the same practical tests. (Unger 1998, pp. 24–25)

We need to maintain a healthy skepticism toward the idea that a specific type of institution—a particular mode of corporate governance, social security system, or labor market legislation, for example—is the only one compatible with a well-functioning market economy.

Market Incentives and Institutions

It is individual initiative that ultimately accounts for all economic progress. The market system is unparalleled in its efficacy in directing individual effort toward the goal of material advancement of society. Early thinking on development policy did not take sufficient account of this. Structuralists downplayed market incentives because they viewed them as ineffective in view of pervasive supply and other “structural” constraints. Socialists downplayed market incentives because they viewed them as inconsistent with the attainment of equity and other social goals.

Both fears have turned out to be groundless. Farmers, entrepreneurs, and investors all over the world, regardless of their income and education levels, have revealed themselves to be quite responsive to price incentives. In Korea and Taiwan, China, the private sector’s strong response to the tax and credit incentives put in place during the early 1960s was a critical instigator of these economies’ growth miracles (Rodrik 1995). In China the dual-track system that allowed farmers to sell their crops in free markets (once their quota obligations were fulfilled) resulted in a large increase in agricultural output and sparked the high growth that has continued to date. After India reformed its cumbersome industrial licensing system, reduced the cost of imported capital goods, and altered relative prices in favor of tradables in the early 1990s, it was rewarded with a sharp increase in investment, exports, and growth. While inequality has gotten worse in some of these economies, poverty levels have been reduced in all of them.

So market incentives work. If this were the entire story, the policy conclusion would be equally straightforward: liberalize all markets as fast as you can. This in fact was the message internalized by the advocates of the Washington consensus and the policymakers who listened to them.

But the experience with development during the past half century reveals another striking fact: the best performing countries were those that liberalized partially and gradually. China, of course, stands out in this respect, as its astonishing success since 1978 is due to a strategy based on dual tracks, gradualism, and experimentation. Save for Hong Kong, which has always been a *laissez-faire* haven, all the East Asian success cases have followed gradualist reform paths. India, which did very well in the 1990s, liberalized only partially. All these countries unleashed the energies of their private sector, but did so in a cautious, controlled manner.

An important reason why gradualist strategies worked in these cases was that they were better tailored to preexisting institutions at home. They therefore economized on institution building (see Qian 2000 for an account of China’s experience along these lines). Korea used a repressed, heavily controlled financial system to channel credit to industrial firms willing to undertake investments. The textbook alternative

of financial liberalization coupled with investment tax credits might have been more efficient on paper, but probably would not have worked as well in the Korea of the 1960s and 1970s nor have paid off so quickly. Rather than relying on dual-track pricing, China could have liberalized agricultural prices completely and then compensated urban dwellers and the treasury through tax reforms, but the new institutions would have taken years, if not decades, to build.

Compare these examples with the wholesale reforms implemented in Latin America and the formerly socialist countries. Because the formerly socialist countries were so radical and borrowed en masse from other countries, their success hinged on the herculean task of creating a wide range of new institutions in short order and, for most of them, from scratch. It is not surprising that the transition has proved more difficult than many economists had expected. The most successful cases, such as Poland, are those that had capitalist institutions that were not entirely destroyed or were recent enough to remember.

Market-oriented reform strategies must recognize not only that institutions matter, but also that altering existing institutions takes time and effort. The need for time and effort is both a constraint—because it implies that first-best price reforms may not be feasible—and an opportunity—because it allows imaginative policy-makers to devise profitable alternatives such as the dual-track system or town and village enterprises in China. Development strategies that overlook possibilities for experimentation and local variation foreclose potentially successful paths for growth.

Implications for International Governance and Conditionality

My argument so far can be summarized in four propositions:

- Market incentives are critical to economic development.
- Market incentives need to be underpinned by strong public institutions.
- Market economies are compatible with a diverse range of institutional arrangements.
- The greater is the fit between market-oriented reforms and preexisting institutional capabilities, the higher is the probability of success.

The first two propositions are now widely accepted, and they form the foundation of an augmented Washington consensus. According to the revised consensus, liberalization, privatization, and global integration are no less important, but they need to be supplemented and supported by reforms in governance. But the importance of the third and fourth propositions is not adequately recognized.

We see the new consensus in operation in a number of areas. In the aftermath of the East Asian crisis, for example, International Monetary Fund (IMF) programs in the region prescribed a long list of structural reforms in business-government relations, banking, corporate governance, bankruptcy laws, labor market institutions, and industrial policy. A key component of the new international financial architecture is a set of codes and standards—on fiscal transparency, monetary and financial policy, banking supervision, data dissemination, corporate governance and struc-

ture, and accounting standards—designed for application in all countries but targeted especially to developing countries. And ever since the Uruguay Round, global trade negotiations have resulted in a number of agreements—on subsidies, intellectual property rights, and investment-related measures—that harmonize practices in the developing countries with those in the more advanced countries.

Thus as the new view of development comes to be operationalized, it results in a ratcheting up of conditionality and a narrowing of the space within which policy can be conducted. In general, this is undesirable for several reasons. First, it is ironic that this is happening at precisely the moment when our comprehension of how the global economy works and what small countries need to do to prosper within it has been revealed to be sorely lacking. It was not so long ago that East Asia's export orientation and high investment rates were assumed to provide protection against the kind of external crisis that periodically rocks Latin America. A common exercise in the aftermath of the 1995 tequila crisis was to compare the two regions' current account deficits, real exchange rates, export-to-GDP ratios, and investment rates to show how East Asia, for the most part, looked "better." East Asia had its critics, of course, but what the critics had in mind was a gradual running out of steam, not the meltdown that transpired. Alan Greenspan was quoted in early 1999 as saying, "I have learned more about how this new international financial system works in the last 12 months than in the previous 20 years" (Friedman 1999, p. 71).

Second, as I have emphasized, market capitalism is compatible with a variety of institutional arrangements. The new consensus either rejects this view (the extreme "convergence" view) or underestimates its significance in practice. The new set of external disciplines comes hand-in-hand with a particular model of economic development that is untested even in the historical experience of today's advanced countries. These disciplines foreclose some development strategies that have worked in the past and others that could work in the future. The narrowing of national autonomy in the formulation of development strategies is a cost for which developing countries are unlikely to receive an adequate reward.

Third, the practical difficulties of implementing many of the institutional reforms under discussion are severely underestimated. Today's developed countries did not get their legal and regulatory institutions overnight. It would be nice if developing countries could somehow acquire developed country institutions, but the safe bet is that this will happen only when they are no longer developing countries. A strategy that tailors market-based reforms to existing institutional capabilities is more likely to bear fruit in the short run.

None of this is to suggest that the specific institutional reforms that dominate the agendas of the Bretton Woods institutions are without merit. No one can seriously oppose the introduction of proper accounting standards or improved prudential supervision of financial intermediaries. While some of the standards are likely to backfire in practice, the more serious concern is that the agendas focus too much on institutional reforms needed to make the world safe for capital flows and therefore necessarily divert political capital and attention from institutional reforms in other areas. The risk is that such an approach privileges freedom of international trade

and capital mobility in the name of “sound” economic policy and that it does so at the cost of neglecting other goals of development policy that might clash with free trade and capital flows.

In the real world, governments face choices: Should they devote education resources to training bank auditors or secondary school teachers? Should they focus on the grand corruption that deters foreign investment or on the petty corruption that most concerns the average household? Should politicians spend their political capital on removing obstacles to international integration or on improving administrative capabilities in the public sector? Should resources be spent on harmonizing legislation with international practice or on public health? Priorities matter when administrative, human, and political capital is scarce. The demands of global integration often imply a different weighting of priorities than that in an appropriately development-focused strategy. Whatever shape the evolving architecture of the international economy takes, therefore, an important goal should be to leave space for developing countries to experiment with their own strategies.

How Important Is International Economic Integration?

The requirements of global economic integration have come to cast a long shadow over the design of development policies. Developing countries are incessantly lectured about the long list of requirements they have to fulfill to integrate into the world economy. The trouble with the current discourse on globalization is that it confuses ends with means. A truly development-oriented strategy requires a shift in emphasis. Integration into the world economy has to be viewed as an instrument for achieving economic growth and development, not as an ultimate goal. Maximizing trade and capital flows is not and should not be the objective of development policy.

No country has developed successfully by turning its back on international trade and long-term capital flows. Very few countries have grown over long periods without experiencing an increase in the share of foreign trade in their national product. As Yamazawa (2000) puts it, “no developing economy can develop within its protected wall” (p. 2). In practice, the most compelling mechanism that links trade with growth in developing countries is that imported capital goods are likely to be significantly cheaper than those manufactured at home. Policies that restrict imports of capital equipment, raise the price of capital goods at home, and thereby reduce real investment levels have to be viewed as undesirable *prima facie*. Exports, in turn, are important because they are needed to purchase imported capital equipment.

But it is equally true that no country has developed simply by opening to foreign trade and investment. The trick in the successful cases has been to combine the opportunities offered by world markets with a domestic investment and institution building strategy to stimulate the animal spirits of domestic entrepreneurs. As mentioned, almost all the outstanding cases have involved partial and gradual opening to imports and foreign investment. Multilateral institutions such as the World Bank, International Monetary Fund, and Organisation for Economic Co-operation and Development regularly give advice predicated on the belief that openness generates

predictable and positive consequences for growth. Yet there is simply no credible evidence that across-the-board trade liberalization is systematically associated with higher growth rates.

The Evidence on Trade Liberalization

Recently Francisco Rodríguez and I (1999) reviewed the extensive empirical literature on the relationship between trade policy and growth.⁶ We reached the conclusion that there is a significant gap between the message that the consumers of this literature have derived and the “facts” that the literature has actually demonstrated. The gap emerges from a number of factors. In many cases the indicators of “openness” used by researchers are problematic as measures of trade barriers or are highly correlated with other sources of poor economic performance. In other cases the empirical strategies used to ascertain the link between trade policy and growth have serious shortcomings, the removal of which results in significantly weaker findings.

Thus the nature of the relationship between trade policy and economic growth remains very much an open question. The issue is far from having been settled on empirical grounds. There are in fact reasons to be skeptical that there is a general, unambiguous relationship between trade openness and growth waiting to be discovered. The relationship is likely to be a contingent one, dependent on a host of country and external characteristics. The fact that practically all of today’s advanced countries embarked on their growth behind tariff barriers, and reduced protection only subsequently, surely offers a clue of sorts. Moreover, the modern theory of endogenous growth yields an ambiguous answer to the question of whether trade liberalization promotes growth. The answer varies depending on whether the forces of comparative advantage push the economy’s resources in the direction of activities that generate long-run growth (through externalities in research and development, expansion of product variety, upgrading of product quality, and so on) or divert them from such activities.

Indeed, the complementarity between market incentives and public institutions that I have repeatedly emphasized has been no less important in trade performance. In East Asia the role of governments in getting exports out during the early stages of growth has been studied and documented extensively (Amsden 1989; Wade 1990). Even in Chile, the exemplar of free market orientation, post-1985 export success has been dependent on a wide range of government policies, including subsidies, tax exemptions, duty drawback schemes, publicly provided market research, and public initiatives fostering scientific expertise. After listing some of the pre- and post-1973 public policies promoting the fruit, fishery, and forestry sectors in Chile, Maloney (1997) concludes that “it is fair to wonder if these, three of the most dynamic export sectors, could have responded to the play of market forces in the manner they have without the earlier and concurrent government support” (pp. 59–60).

The appropriate conclusion to draw from all this is not that trade protection should be preferred to trade liberalization as a rule. There is no evidence from the past 50 years that trade protection is systematically associated with faster growth.

The point is simply that the benefits of trade openness should not be oversold. When other worthwhile policy objectives compete for scarce administrative resources and political capital, deep trade liberalization often does not deserve the high priority it typically receives in development strategies. This is a lesson of particular importance to countries in the early stages of reform, such as those in Africa.

The Evidence on Capital Account Liberalization

The evidence on the benefits of capital account liberalization is even weaker.⁷ On paper, the appeal of capital mobility is obvious. In the absence of market imperfections, freedom to trade enhances efficiency, and that is as true of trade in paper assets as it is of trade in widgets. But financial markets suffer from various syndromes—myopia, information asymmetries, agency problems, self-fulfilling expectations, and bubbles (rational and otherwise)—to an extent that makes their economic analysis inherently a second-best one. No amount of institutional tinkering is likely to make a significant difference to that basic fact of life.

The question of whether developing nations should be pushed to open their capital accounts (in an orderly and progressive manner, as is now recommended by the IMF) can ultimately be resolved only on the basis of empirical evidence. While there is plenty of evidence that financial crash often follows financial liberalization (see Williamson and Mahar 1998 for a survey), there is very little evidence that higher rates of economic growth follow capital account liberalization. Quinn (1997) reports a positive association between capital account liberalization and long-run growth, while Grilli and Milesi-Ferretti (1995), Rodrik (1998a), and Kraay (1998)—the last author using Quinn's (1997) indicator of capital account restrictions—find no relationship. Klein and Olivei (1999) report a positive relationship, but one driven largely by the experience of the developed countries in their sample. This is a field of inquiry that remains in its infancy, and there is clearly much more to be learned. The most that can be said at present is that convincing evidence on the benefits of capital account liberalization has yet to be produced.

Among all the arguments in favor of international capital mobility, perhaps the most appealing one is that such mobility serves a useful disciplining function for government policy. Governments that have to be responsive to investors cannot squander their society's resources as easily. As Summers (1998) puts it, "market discipline is the best means the world has found to ensure that capital is well used."

The idea is attractive, but once again one has to question its empirical relevance. When foreign creditors suffer from the syndromes noted above, a government intent on irresponsible spending finds it easier to finance its expenditures when it can borrow from abroad. Moreover, for such a government even domestic borrowing becomes politically less costly because in a world of free capital mobility there is no crowding out of private investors (since they can borrow from abroad). In both instances international financial markets allow reckless spending that might not have taken place in their absence. Conversely, the discipline that markets exert in the aftermath of crises can be excessive and arbitrary, as discussed previously. As Willett

(1998) points out, the appropriate characterization of market discipline is that it comes too late, and that when it comes it is typically too much.

A recent paper by Mukand (1998) develops the analytics of such situations nicely. Consider the following stylized setup suggested by Mukand's framework. Let there be two actors, a government (G) and a foreign investor (F), which have to decide what actions to pursue when the underlying state of the world is not observable. The state of the world can be either neat or messy. G receives a private signal about the state of the world and then chooses a policy, which is then observed by F. The policy can be either orthodox or heterodox. Assume that the orthodox policy produces a larger surplus in aggregate when the state of the world is neat, and the heterodox policy a larger surplus when the state of the world is messy. F wants to invest only when the policy and the expected state match (orthodox and neat or heterodox and messy). In addition, F believes (perhaps incorrectly) that the productivity of its investment will be higher in the orthodox-neat scenario than in the heterodox-messy one and will invest more when it expects the first scenario.

Mukand demonstrates that the government may have two reasons to follow the orthodox policy under these circumstances, even when it receives a signal that the underlying state of the world is messy (and therefore the heterodox policy would have been more appropriate). He calls the resulting biases "conformity bias" and "good-news bias":

Conformity bias: Let F have a strong and unmovable prior belief that the state of the world is neat. Even if G's posterior belief that the state is messy is strong, G may want to follow orthodox policy anyway because it will not be able to sway F's (posterior) belief and may be better off having the investment and following the wrong policy than not having the investment and following the right (aggregate-surplus-maximizing) policy.

Good-news bias: When F's posterior belief can be affected by G's choice of policy, G may want to follow orthodox policy to signal a neat state, because more investment will be forthcoming when F expects the neat state rather than the messy one (assuming that there is a match between the expected state of the world and policy in both cases).

For the good-news scenario to materialize, it is not necessary for the productivity of investment to be higher in the orthodox-neat scenario than in the heterodox-messy one. All that is needed is that the foreign investor believe so. In either case the government finds itself driven by "market sentiment" to follow policies that are inappropriate and fall short of the optimum.

Governments do need discipline, of course. However, in modern societies this discipline is provided by democratic institutions—elections, opposition parties, independent courts, parliamentary debate, a free press, and other civil liberties. Governments that mess up the economy are punished at the polls. Broad cross-national evidence suggests that democratic nations tend to be pretty good at maintaining responsible fiscal and monetary policies. Most significant cases of fiscal profligacy occur under authoritarian regimes rather than democratic ones. It was

military dictatorships that got Latin America into its debt crisis, and democracies that cleaned up the mess. In Asia democratic countries such as India and Sri Lanka have exemplary macroeconomic records by African or Latin American standards. Africa's only long-running democracies (Botswana and Mauritius) have done an excellent job of managing booms and busts in the prices of their main exports (diamonds and sugar). Among the transition economies, the most successful stabilizations have occurred in the most democratic countries. There is a strong negative association between the Freedom House index of democracy and the average inflation rate in a sample of more than 100 countries, after controlling for per capita income. The international-capital-mobility-as-discipline position embodies a view of politics that is at best partial, and at worst harmful to democracy.

Finally, the pursuit of the capital account liberalization agenda has the effect of crowding out policymakers' agenda and diverting their energies from national development efforts. A finance minister occupied with mollifying investor sentiment and marketing the economy to foreign bankers can spend no time on traditional development concerns: reducing poverty, mobilizing resources, and setting investment priorities. Global markets, not domestic priorities, end up dictating policy.

Conclusion

The lesson of the 20th century is that successful development requires markets underpinned by solid public institutions. Today's advanced industrial countries—Japan, the United States, Western European nations—owe their success to having evolved their own workable models of a mixed economy. While these societies are alike in the emphasis they place on private property, sound money, and the rule of law, they are dissimilar in many other areas: their practices in labor market relations, social insurance, corporate governance, product market regulation, and taxation differ substantially.

All these models are in constant evolution, and none is without its problems. European-style welfare capitalism seemed especially appealing in the 1970s. Japan became the model to emulate during the 1980s. And the 1990s were clearly the decade of U.S.-style freewheeling capitalism. Evaluated in an appropriately historical perspective, all these models have been equally successful. The evidence from the second half of the 20th century is that none of these models clearly dominates the others. It would be a mistake to hold up U.S.-style capitalism as the model toward which the rest of the world must converge.

Of course, all successful societies are open to learning, especially from useful precedents in other societies. Japan is a good example of this. When Japan reformed and codified its legal system under the Meiji restoration, it used Germany's civil and commercial law as the primary model. So my emphasis on institutional diversity and nonconvergence should not be viewed as a rejection of institutional innovation through imitation. What is important is that imported "blueprints" be filtered through local practices and needs. Once again, Japan provides the example. As Berkowitz, Pistor, and Richard (1999) discuss, Japan's selection of the German legal system was an informed choice, not an imposition from

abroad: “extensive debates about the adoption of English or French law, and several drafts based on the French model, preceded the promulgation of codes that were largely based on the German model” (p. 11). In other words, Japanese reformers consciously selected from the codes that were available those that seemed most suited to their circumstances.

What is true of today’s advanced countries is also true of developing countries. Economic development ultimately derives from a homegrown strategy, not from the world market. Policymakers in developing countries should avoid fads, put globalization in perspective, and focus on domestic institution building. They should have more confidence in themselves and in domestic institution building, and place less faith in the global economy and the blueprints that emanate from it.

Notes

1. This section draws on Rodrik (1999a, chap. 4).
2. This point is disputed by many and goes against the official view of the International Monetary Fund (Fischer 1998). The argument that “structural” aspects of the East Asian model were not at the root of the crisis is put well by Stiglitz (1999) and Radelet and Sachs (2000). This is not to say that these economies did not have structural weaknesses, in particular, an overreliance on government steering of the economy that had probably outlived its usefulness. But as Stiglitz points out, financial crises break out with some regularity in economies ranging from those of Scandinavia to that of the United States, with very different types of economic management and standards of transparency.
3. This section borrows heavily from Rodrik (forthcoming).
4. Glaeser, Johnson, and Shleifer (1999) attribute the more impressive development of equity markets in Poland compared with those in the Czech Republic to Poland’s stronger regulations upholding minority shareholder rights and guarding against fraud.
5. Perhaps Europe will come back into fashion. The *New York Times* recently published a major feature article with the title “Sweden, the Welfare State, Basks in a New Prosperity” (Andrews 1999).
6. Our detailed analysis covers the five papers that are probably the best known in the field: Dollar (1992), Sachs and Warner (1995), Ben-David (1993), Edwards (1998), and Frankel and Romer (1999).
7. This discussion on capital account convertibility is based on Rodrik (2000).

References

- Alesina, Alberto, and Allan Drazen. 1991. “Why Are Stabilizations Delayed?” *American Economic Review* 81 (5): 1170–88.
- Amsden, Alice H. 1989. *Asia’s Next Giant: South Korea and Late Industrialization*. New York: Oxford University Press.
- Andrews, Edmund L. 1999. “Sweden, the Welfare State, Basks in a New Prosperity.” *New York Times*, 8 October 1999, Business.
- Ben-David, Dan. 1993. “Equalizing Exchange: Trade Liberalization and Income Convergence.” *Quarterly Journal of Economics* 108 (3): 653–79.
- Berkowitz, Daniel, Katharina Pistor, and Jean-François Richard. 1999. “Economic Development, Legality and the Transplant Effect.” Harvard University, Kennedy School of Government, Cambridge, Mass.

- Bordo, Michael D., Claudia Goldin, and Eugene N. White, eds. 1998. *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century*. Chicago: University of Chicago Press.
- Collins, Susan, and Barry Bosworth. 1996. "Economic Growth in East Asia: Accumulation versus Assimilation." *Brookings Papers on Economic Activity* (2): 135–91.
- Dollar, David. 1992. "Outward-Oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976–85." *Economic Development and Cultural Change* 40: 523–44.
- Edwards, Sebastian. 1998. "Openness, Productivity and Growth: What Do We Really Know?" *Economic Journal* 108: 383–98.
- Fischer, Stanley. 1998. "In Defense of the IMF." *Foreign Affairs* 77 (4): 103–06.
- Frankel, Jeffrey, and David Romer. 1999. "Does Trade Cause Growth?" *American Economic Review* 89 (3): 379–99.
- Friedman, Thomas L. 1999. "A Manifesto for the Fast World." *New York Times Magazine*, 28 March, 40–61, 70–71, 84, 97–97.
- Glaeser, Edward, Simon Johnson, and Andrei Shleifer. forthcoming. "Coase v. the Coasians: The Regulation and Development of Securities Markets in Poland and the Czech Republic." *Quarterly Journal of Economics*.
- Grilli, Vittorio, and Gian Maria Milesi-Ferretti. 1995. "Economic Effects and Structural Determinants of Capital Controls." *IMF Staff Papers* 42 (3): 517–51.
- Hausmann, Ricardo, and Michael Gavin. 1996. "Securing Stability and Growth in a Shock-Prone Region: The Policy Challenge for Latin America." Inter-American Development Bank, Washington, D.C.
- Hoff, Karla, and Joseph E. Stiglitz. 2001. "Modern Economic Theory and Development." In Gerald Meier and Joseph E. Stiglitz, eds., *The Future of Development Economics in Perspective*, Oxford: Oxford University Press.
- Jacoby, Sanford M. 1998. "Risk and the Labor Market: Societal Past as Economic Prologue." University of California at Los Angeles, Institute of Industrial Relations.
- Klein, Michael, and Giovanni Olivei. 1999. "Capital Account Liberalization, Financial Deepening, and Economic Growth." NBER Working Paper 7384. National Bureau of Economic Research, Cambridge, Mass.
- Kraay, Aart. 1998. "In Search of the Macroeconomic Effects of Capital Account Liberalization." World Bank, Macroeconomics and Growth, Washington, D.C.
- Maloney, William F. 1997. "Chile." In Laura Randall, ed., *The Political Economy of Latin America in the Postwar Period*. Austin, Tex.: University of Texas Press.
- Mukand, Sharun W. 1998. "Globalization and the 'Confidence Game': Policy Making in the Open Economy." Tufts University, Department of Economics, Medford, Mass.
- North, Douglass C., and R. Thomas. 1973. *The Rise of the Western World: A New Economic History*. Cambridge: Cambridge University Press.
- North, Douglass C., and Barry Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *Journal of Economic History* 49: 803–32.
- Qian, Yingyi. 2000. "Institutional Foundations of China's Market Transition." In Boris Pleskovic and Joseph E. Stiglitz, eds., *Annual World Bank Conference on Development Economics 1999*. Washington, D.C.: World Bank.
- Quinn, Dennis. 1997. "The Correlates of Change in International Financial Regulation." *American Political Science Review* 91 (3): 531–51.
- Radelet, Steven, and Jeffrey Sachs. 2000. "The Onset of the Asian Financial Crisis." In Paul Krugman, ed., *Currency Crises*. Chicago: University of Chicago Press.

- Rodríguez, Francisco, and Dani Rodrik. 1999. "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence." NBER Working Paper 7081. National Bureau of Economic Research, Cambridge, Mass. (Also forthcoming in Ben Bernanke and Kenneth S. Rogoff, eds., *NBER Macroeconomics Annual, 2000*. Cambridge, Mass.: National Bureau of Economic Research.)
- Rodrik, Dani. 1995. "Getting Interventions Right: How South Korea and Taiwan Grew Rich." *Economic Policy* 20: 55–107.
- . 1998a. "Who Needs Capital-Account Convertibility?" In Stanley Fischer and others, eds., *Should the IMF Pursue Capital-Account Convertibility? Essays in International Finance* 207. Princeton University, Department of Economics, International Finance Section, Princeton, N.J.
- . 1998b. "Why Do More Open Economies Have Bigger Governments?" *Journal of Political Economy* 106 (5): 997–1032.
- . 1999a. *The New Global Economy and Developing Countries: Making Openness Work*. Washington, D.C.: Overseas Development Council.
- . 1999b. "Where Did All the Growth Go? External Shocks, Social Conflict, and Growth Collapses." *Journal of Economic Growth* 4 (December 1999): 385–412.
- . 2000. "Governing the World Economy: Does One Architectural Style Fit All?" In Susan M. Collins, ed., *Brookings Trade Forum: 1999*. Washington, D.C.: Brookings Institution Press.
- . Forthcoming. "Institutions for High-Quality Growth: What They Are and How to Acquire Them." *Studies in Comparative International Development*.
- Sachs, Jeffrey, and Andrew Warner. 1995. "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity* (1): 1–118.
- Stiglitz, Joseph E. 1999. "Knowledge and Development: Economic Science, Economic Policy, and Economic Advice." In Boris Pleskovic and Joseph E. Stiglitz, eds., *Annual World Bank Conference on Development Economics 1998*. Washington, D.C.: World Bank.
- Summers, Lawrence H. 1998. "Building an International Financial Architecture for the 21st Century." Remarks to the Cato Institute, 16th Annual Monetary Conference, Washington, D.C., 22 October. [www.usconsulate.org.hk/gf/1998/1022b.htm].
- Tornell, Aaron, and Philip R. Lane. 1999. "The Voracity Effect." *American Economic Review* 89 (1): 22–46.
- Unger, Roberto Mangabeira. 1998. *Democracy Realized: The Progressive Alternative*. New York: Verso.
- Wade, Robert. 1990. *Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization*. Princeton, N.J.: Princeton University Press.
- Willett, Thomas D. 1998. "International Financial Markets as Sources of Crises or Discipline: The Too Much, Too Late Hypothesis." Claremont McKenna College, Department of Economics, Claremont, Calif.
- Williamson, John, and Molly Mahar. 1998. *A Survey of Financial Liberalization*. Essays in International Finance 211. Princeton University, Department of Economics, International Finance Section, Princeton, N.J.
- World Bank. 1997. *World Development Report 1997: The State in a Changing World*. New York: Oxford University Press.
- Yamazawa, Ipeei. 2000. "Regional Cooperation in a Changing Global Environment: Success and Failure of East Asia." Paper prepared for UNCTAD X High-Level Round-Table on Trade and Development, Bangkok, 12 February 2000.

Comment on “Consensus Building, Knowledge, and Conditionality,” by Paul Collier, and “Development Strategies for the 21st Century,” by Dani Rodrik

Nancy Birdsall

In these remarks I propose that development strategies for the 21st century cannot continue to overlook the middle class. I begin with comments linking global governance to a new style of development assistance. This new style, of which Paul Collier and Dani Rodrik are leading proponents, is based on a revisionist view of the fundamental development challenge. This revisionist view sees homegrown institutions as the scarce resource most critical to countries’ development success in a global market. In the advanced economies with democratic systems it is the middle class that supports, and is supported by, these homegrown, market-sustaining institutions. But who in fact is (or could be) the middle class in developing countries? I provide some information on middle-income households in developing countries and suggest some implications for development policy.

Global Governance and Imperfect Representation

The demonstrations in Washington, D.C., during the April 2000 meetings of the International Monetary Fund (IMF) and the World Bank were fundamentally about global governance. They reminded us that the Bretton Woods and other international financial institutions are caught in a squeeze. On one side are those who insist they become supranational vehicles for the good and the just—be it labor standards, a cleaner environment, or uncorrupted and more transparent governments. This camp includes people who would have these global institutions become the major force for more efficient and stable global markets—“making the world safe for capital flows” is Rodrik’s apt characterization—for example, by enforcing disclosure of national financial and macroeconomic indicators. The demonstrators and IMF management are on the same side of this question. On the other side are Collier, Rodrik, and former World Bank Chief Economist Joseph Stiglitz (1998), who decry the leverage of these institutions over policy decisions in independent nations, the breakdown of national auton-

Nancy Birdsall is senior associate at the Carnegie Endowment for International Peace in Washington, D.C.
Annual World Bank Conference on Development Economics 2000
©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

omy implied by their leverage, and their lack of adequate accountability to the citizens of countries affected by their decisions.

The IMF and the World Bank are global institutions with inevitably imperfect representation. Borrowing member governments most affected by the institutions' policies have limited voice. And many member governments do not adequately represent the interests of their own citizens. (Even in mature democracies citizens cannot be represented well on any specific issue.) Transparency, disclosure, and the involvement and pressures of civil society groups in member countries are partial—but only partial—solutions to these problems of imperfect representation.

Collier and Rodrik are leading proponents of new thinking about the business of development assistance *given* this system in which power and money are poorly aligned with representation of and accountability to affected parties. Both assume as a starting point that development assistance will be more effective where it eschews policy leverage ("conditionality") and instead concentrates on supporting national autonomy, local institutions, and the construction of a sustainable local consensus. This is because, as Rodrik argues, the key to past development success, once the worst macroeconomic imbalances were corrected, was not any particular recipe—certainly not a particular trade or industrial strategy, and not necessarily or primarily integration into world trade and capital markets. The key was effective and homegrown market-sustaining public institutions able to manage social conflicts, the constant demands of changing markets, and openness itself—and as Collier puts it, able to build and sustain a political consensus around policy reform.

But where will the market-sustaining institutions for managing change and building consensus come from? Beyond a hands-off tolerance for diversity and national autonomy, what principles might guide development assistance efforts to support such institutions? In the advanced Western economies the middle class has been the backbone of market-sustaining institutions. Might a comparable middle class play that role in developing countries?

The Middle Stratum in Developing Countries: New Opportunities and New Anxieties

The development literature and economists in general have been virtually silent on the middle class.¹ Economists have no simple, cross-societal, time-invariant definition of the middle class. But we can compare across countries the absolute and relative positions of the middle *stratum* (Nelson 2000)—for example, households with per capita income between 75 and 125 percent of the median—recognizing that in many countries this group may not fit our prior notion of the middle class.²

Middle-stratum households in developing countries are obviously much poorer in absolute terms than their middle-income counterparts in rich countries. They also tend to be poorer relative to their richer fellow citizens. Even in such emerging markets as Brazil and Poland, middle-stratum households have only about a tenth of the income (in purchasing power parity terms) of their counterparts in the United States. They are also much less educated. In Costa Rica adults in these households

have only about 6 years of education, and in Brazil about 4 years, compared with almost 13 years in the United States. The difference in education between the richest 5 percent of households and middle-income households is also much greater than that in the United States. Costa Rican adults in the richest households have an average of about 13 years of schooling—7 years more than adults in middle-income households. In the United States the difference is just 2 years: 15 years compared with 13.

Middle-income households in developing countries are not only closer to poverty (obviously, given their lower average income), but also probably more downwardly mobile. As many as 30 percent of nonpoor households in Indonesia and 22 percent in Peru are likely to become poor in the next three to six years, based on three-year fluctuations in reported per capita household expenditures. Of households in Peru sampled in 1990, almost half of those initially in the middle quintile of the income distribution had slipped down to the two lower quintiles by 1996; in the United States over a recent longer period, only a third of households suffered downward mobility.³

Surveys find the middle stratum in many developing countries an anxious and unhappy group—even where its members have prospered. For example, Graham and Pettinato (2000) show that many households in Peru and Russia that enjoyed rapid income growth in the past actually report that they are worse off, including in income terms. This may be the case in the richest and most advanced economies as well.

Anxiety is in part the outcome of the greater uncertainty of more market-driven economies operating in a more integrated and volatile global economy. In Latin America the return to growth in the 1990s brought modest income gains on average but more economic insecurity for middle-stratum households. Growth in private sector employment did not make up for the loss of secure jobs in government and in state-owned enterprises (except in Mexico and Central America, where growth in the *maquila* sector has been substantial). Even those who kept secure public sector jobs lost ground in terms of wage growth compared with private sector and informal sector workers, as in Peru (Birdsall, Graham, and Pettinato 2000, based on Saavedra 1998).

An increase in “unprotected” jobs with no written contract or social benefits for workers probably affected the middle stratum most because the truly poor never enjoyed protected jobs in the first place. Rodrik (1999) provides evidence suggesting that not only jobs, but wages became less secure as the volatility of the real average wage increased in the 1990s, driven mainly by macroeconomic shocks associated with more open capital markets.

In the formerly socialist economies of Europe and Central Asia middle-stratum households were hit hard by the transition. Even in Poland, where average annual per capita growth was positive between 1986 and 1995, the number of households in this category shrank more than 15 percent, and their income share shrank proportionately. In Hungary between 1991 and 1994, the number of households in this group shrank by 13 percent, and their income share by 17 percent.

In East Asia the financial crisis hit an emergent but still vulnerable middle-income group. Declining employment and falling wages in construction and manufacturing

hurt urban salaried workers, and heavy reliance on high interest rates to defend falling currencies hurt small business owners. Poor rural households and urban households in the low-productivity informal sector were obviously vulnerable to the economywide recessions. But a second group of people—those who might be called market-friendly urban strivers (Birdsall and Haggard 2000)—found themselves vulnerable for the first time as a result of the crisis. These people belong to households at or below incomes of about \$5,000 (in purchasing power parity terms). In Indonesia this group is in the top 40 percent of the income distribution, in Thailand closer to the middle, and in the Republic of Korea closer to the bottom, although still mostly above the poverty line. Emergency measures were designed to provide a safety net for the poor but not particularly for this second group.

Easterlin (1995) has pointed out that people with low income care most about absolute gains in income, while people with higher income care more about their income relative to the incomes of those better off. If relative income matters, then where income and wealth are growing fastest for those at the top (for example, for the highly educated, as returns to education increase worldwide), households around the median are likely to suffer "middle-income stress." This will be especially true in developing countries, because of the bigger gaps in education between the rich and the middle class. Moreover, globalization of consumption standards might well be exacerbating such stress (Nike sales, for example, grew by 82 percent in Asia and 91 percent in Latin America in one year in the mid-1990s). Stress in middle-income households is low (but increasing) in Eastern Europe, and high (but declining) in Latin America (table 1).

The point is not to bemoan the effects of market reforms and globalization on the middle stratum or emerging middle class—many of those now in the middle stratum are there because market reforms generated new opportunities to escape poverty. Instead, it is to underline how little consideration we have given to what determines the size and income share of the middle stratum, how the middle stratum affects

Table 1. Size and Stress of the Middle Stratum in Selected Countries, Various Years, 1990s

Country	Year	Population share of middle stratum (percent)	Income share middle stratum (percent)	Middle class stress
Sweden	1995	38.0	31.6	1.43
United Kingdom	1995	33.1	26.0	2.07
United States	1999	24.2	17.6	2.67
Hungary	1994	43.8	35.0	1.77
Poland	1995	36.0	31.6	1.75
Slovak Republic	1992	58.2	54.6	1.30
Brazil	1996	20.7	9.6	7.32
Costa Rica	1997	24.5	17.6	2.99
Peru	1997	21.4	13.1	4.14

Note: The middle stratum is defined as households with per capita income between 75 and 125 percent of the median per capita income of all households. Middle-class stress is defined as the ratio of the median income of the households commanding 50 percent of total national income to the median income of all households.

Source: Birdsall, Graham, and Pettinato 2000.

market reforms and the integration of economies into global markets, and what role the members of this group play in the political and economic discourse and institutional underpinnings of the policies in their countries.

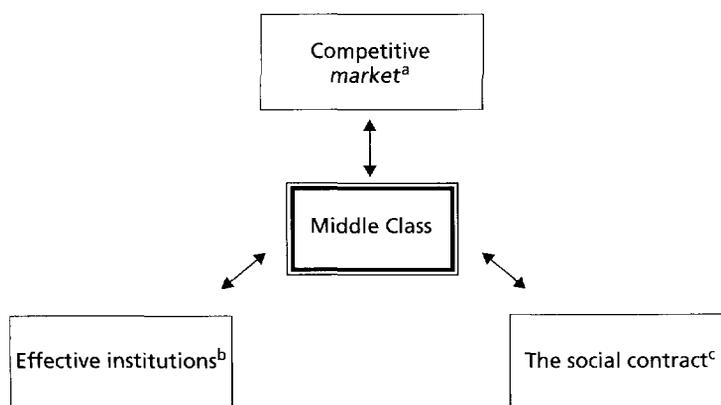
Nurturing a Market-Friendly Middle Class

Is a market-friendly middle class relevant to development policy and to the conduct of the development assistance business? What are the causes and consequences of a market-friendly middle class?

Some market reforms—privatization, tax reform, trade and financial sector liberalization—have not been easy on middle-income households. But a second round of *institution-building reforms* could greatly increase the stake of the emerging middle class in the market system. Better banking, better schools, property rights, the rule of law, contract enforcement, and a labor market organized around collective bargaining can all be thought of as contributing to a market-sustaining middle class. Reforms could be structured in a way to exploit this positive potential.

The process probably also works the other way around: an independent middle class will likely contribute to the healthy functioning of those same institutions (figure 1). Rodrik's emphasis on managing social conflict and Collier's on moving beyond conditionality to consensus are healthy reminders that some decisions in market economies are best made politically. Some economic reforms—relating to labor, pensions, and schools, for example—are more likely to be sustained if they emerge from a political process and not solely a technical analysis. An informed, educated

Figure 1. The Market-Friendly Middle Class



a. The focus of the first generation of economic reforms (stabilization, privatization, trade and financial sector liberalization, and so on).

b. The focus of the second generation of reforms (property rights, rule of law, judicial reform, education reform, and so on).

c. An outcome of representative and participatory democracy.

middle class can provide considerable institutional ballast in the volatile world of democratic politics. Indeed, Aristotle pointed out that the middle class is the natural ally of democratic institutions. The role of the middle class seems particularly important in settings where ethnic, racial, cultural, religious, and linguistic differences put a premium on effective and independent regulatory and judicial institutions.

Experience in the West suggests that the middle class is at the heart of any effective social contract. To be sustained politically, the social safety net programs built into a social contract must reflect the outcome of a political bargain that meets the needs of the middle class as well as the poor. In developing countries the poor might benefit more from a small portion of massive, middle-class-driven social insurance programs than from a large portion of targeted services; this has been the case in Europe and the United States (Nelson 2000, referring to Skocpol 1991; Goodin and LeGrand 1987). In open economies subject to global shocks, targeting the poor with special transfers may—if it implies a new burden on an insecure middle class—be less politically acceptable and sustainable than more comprehensive social insurance programs. A social contract, moreover, extends far beyond the narrow notion of the safety net; it includes the political freedoms and rights of participation governing a broad range of social and economic decisions. The ability to design and maintain this broad social contract may ultimately depend on the allegiance and intelligence of a stakeholding middle class.

It is true that the middle class is endogenous, an outcome of as well as an input to healthy institutions (see figure 1). My aim is not to propose an analytic framework or a new magic bullet for development, but simply to argue that we would do well to begin incorporating the idea of the middle class into our thinking about development. As Joe Stiglitz noted in his keynote address, ideas can be important in changing policy.

Finally, in focusing on the middle class we are not neglecting the poor. In the end, the sustainable growth on which poverty reduction depends requires adequate institutions. History tells us that the middle class is the bedrock of those institutions, at least in the democratic and open market systems to which so many people in the developing world justifiably aspire. Poverty reduction may ultimately require a market-friendly middle class.

Notes

1. A search of titles on the World Bank Group Web site yielded two hits on *middle class*—both referring to the same article by William Easterly—and “maximum number of hits/overload” (apparently implying hits beyond five digits) on *poverty*. A search covering titles and abstracts on the site “EconLit (1969–2000)” yielded 218 hits on *middle class* or *middle strata* and almost 10,000 on *poverty* or *poor*.

2. Data sources for this and all other statements about the middle stratum are cited in Birdsall, Graham, and Pettinato (2000).

3. For Indonesia, see Pritchett, Suryahadi, and Sumarto (2000), whose estimates are based on three-year fluctuations in reported per capita household expenditures. For Peru, see Birdsall, Graham, and Pettinato (2000) for data on the transition of households in the middle quintile in 1990 into the lower two quintiles in 1996.

References

- Aghion, Philippe, Eve Caroli, and Cecilia García-Peñalosa. 1999. "Inequality and Economic Growth: The Perspective of the New Growth Theories." *Journal of Economic Literature* 37 (4): 1615–60.
- Birdsall, Nancy, and Stephan Haggard. 2000. *After the Crisis: The Social Contract and the Middle Class in East Asia*. Washington, D.C.: Carnegie Endowment for International Peace.
- Birdsall, Nancy, Carol Graham, and Stefano Pettinato. 2000. "Stuck in the Tunnel: Have New Markets Muddled the Middle Class?" Carnegie Endowment for International Peace, Washington, D.C.
- Easterlin, Richard A. 1995. "Will Raising the Incomes of All Increase the Happiness of All?" *Journal of Economic Behavior and Organization* (Netherlands) 27 (1): 35–47.
- Goodin, Robert E., and Julian LeGrand. 1987. *Not Only the Poor: The Middle Classes and the Welfare State*. London: Allen and Unwin.
- Graham, Carol, and Stefano Pettinato. 2000. "Hardship and Happiness: Mobility and Public Perceptions during Market Reforms." *World Economics*, Vol.1, No.4.
- Nelson, Joan. 2000. "The Politics of Social Sector Reforms." Overseas Development Council, Washington, D.C.
- Pritchett, Lant, Asep Suryahadi, and Sudarno Sumarto. 2000. "Quantifying Vulnerability to Poverty: A Proposed Measure with Application to Indonesia." World Bank, Environment and Social Development Sector Unit, Washington, D.C.
- Rodrik, Dani. 1999. "Why Is There So Much Economic Insecurity in Latin America?" Harvard University, John F. Kennedy School of Government, Cambridge, Mass.
- Saavedra, Jaime. 1998. "Empleo, Productividad e Ingresos. Peru (1990-1996)." Equipo Técnico Países Andinos. International Labor Organization.
- Skocpol, Theda. 1991. "Universal Appeal: Politically Viable Policies to Combat Poverty." *Brookings Review*, 9:1, summer: 29–33.
- Stiglitz, Joseph E. 1998. "More Instruments and Broader Goals: Moving Towards the Post-Washington Consensus." Wider Annual Lectures 2. Wider: The U.N. University, (January).

Comment on "Consensus Building, Knowledge, and Conditionality," by Paul Collier, and "Development Strategies for the 21st Century," by Dani Rodrik

François Bourguignon

Paul Collier and Dani Rodrik make concrete the renewal that seems to be taking place in the way we think about economic development, development policy, and the fundamental role of international aid and international development institutions. Interestingly, both echo the mounting worry about what globalization may imply for development and development policies in poor countries. Unlike the demonstrators whose slogans were heard in the Washington streets during the IMF and World Bank spring meetings, however, they contribute genuinely new ideas to the debate and point to promising directions for future action and research in development.

The message seems simple. Development thinking and development policy must include the role of institutions, and possibly changes in institutions, in the design and analysis of policy reforms. Largely because not enough attention was paid to institutions in the past, policy reforms have been less effective than expected, and since the 1970s many developing countries have grown more slowly than expected. For the same reason the so-called transition in the 1990s proved much harder in some countries than expected. And strict loan conditions requiring policy reforms without due consideration of existing institutions or the changes required in them made aid less effective than it could have been in relieving world poverty.

It is always encouraging to be told that although what was done or thought in the past was wrong, we now know why and can think and act correctly. Even though I share with Collier and Rodrik the conviction that institutions matter very much for development, however, we still need to define with some precision what this means for development policy and development analysis.

Rodrik's account of the development experience of the past two or three decades and what he believes were errors of appreciation by the international development community is both convincing and impressive. It is perfectly clear from this account that too much emphasis was put on specific policy reforms without due regard for the institutional framework in which those reforms had to be implemented. But it is much

François Bourguignon is editor of the *World Bank Economic Review*.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

less clear what the fit between institutions and reforms should be or how institutions should be modified. Is it sufficient, for example, to say that democracy or social insurance institutions facilitate policy reforms? How is a society made more democratic, and how can it be persuaded to put social insurance institutions in place?

Collier's critique of the "aid for reform" doctrine since the mid-1970s is equally impressive and convincing. But his answer to the preceding questions may be less so. In particular, why would it be easier to build a consensus about the role of institutions and the effects of changing them than it was to build a consensus earlier about policy reforms?

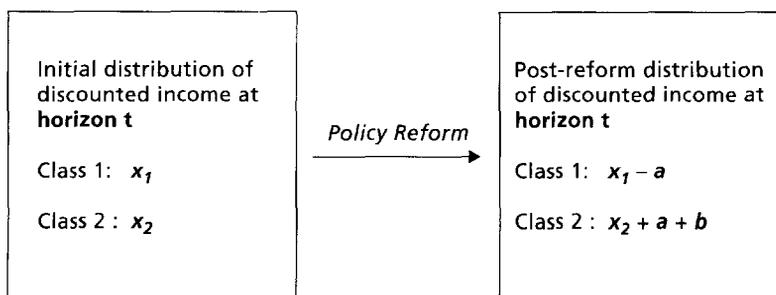
I have no better answer to these difficult questions. But I have learned that we can often make progress in this kind of debate by looking at the questions from a different angle or in a different framework. This is what I intend to do. I first try to formalize the issue of the fit between institutions and policy reform in a simple framework. I then look at the problem of institutional change that runs through the articles by Collier and Rodrik. Because it would be difficult to generalize in dealing with such difficult issues, I concentrate on institutions concerned with redistribution. We shall see, however, that this subject covers quite a bit of ground.

A simple model will prove useful for the discussion (figure 1). Assume that society comprises two classes of individuals, so that the state of the economy may be represented by two numbers that stand for the welfare levels of the classes. Welfare is defined over a given time horizon and takes into account all kinds of uncertainty. Without the reform the two numbers would be x_1 and x_2 . With the reform—for example, lowering tariffs and nontariff barriers—the numbers would be $x_1 - a$ and $x_2 + a + b$, where a and b are numbers that represent differentiated changes in welfare.

Consider two cases. The first is the familiar Pareto dominance case, in which a is negative and $a + b$ is positive. Presumably this case is the less interesting one because the reform should raise no debate and should be undertaken with consensus. In the second case the overall gain is still b , but a is now positive. Thus, the gain is concentrated in the hands of class 2, while class 1 loses a .

If it were possible for class 2 to redistribute a or a little more to class 1, the reform would be undertaken because a "consensus" would be reached. This is the old com-

Figure 1. The Structure of Decisionmaking in Policy Reform



pensation principle used in welfare theory. The problem arises when this redistribution is impossible, either because the instruments are not available—for example, the redistribution authority has no way to unambiguously identify the people in class 1—or because the benefits of the reform will materialize only in the future and class 2 cannot effectively commit itself to redistributing them when they show up. There may also be a way to redistribute a from class 2 to class 1 but at a cost—for example, distortions—that exceeds b . Whether or not the reform is undertaken will then depend on political economy or public choice factors. Aid for reform could be one of those factors or simply an additional element in the political economy of the decision, as Collier explains very well.

This formalization is elementary, and the framework could be made more realistic in many ways. As simple as it is, however, it illustrates some important points in the debate on development policy and policy reforms:

- The values of a and b that are associated with the reform—or that may be functions of the characteristics of the reform—may be imperfectly known (the “knowledge” part of the title of Collier’s article). Clearly, ambiguity about these values necessarily makes the decision about the policy reform more difficult. But some divergence in the estimates of a and b by economic experts, policymakers, and people in classes 1 and 2 may be justified. For example, the rates of time preference or risk aversion may differ across these agents. As mentioned, a may even be negative and $a + b$ positive—corresponding to a Pareto improvement—according to the objective function of international development institutions, but a may be positive, and thus a loss, from the point of view of class 1.
- The difference between aggregated and disaggregated benefits should be emphasized. To an external observer, the positive aggregate benefit b would seem to be a good argument for undertaking the reform. But if there is no way to compensate class 1 so that it can gain from the reform, the reform may be blocked, depending on the kind of public decision process used.
- Because of the difference between aggregated and disaggregated benefits, the effectiveness of existing redistribution channels appears to be a key factor in whether or not the “collectively profitable” reform will pass.

Trade liberalization is a good illustration of these points. Some literature emphasizes the problem of distributing the gains from trade (or from inflows of foreign capital) across the population. Some theoretical models show that where lump-sum transfers are impossible, the aggregate gains from trade might be dissipated in redistribution channels (Feenstra 1987; Lewis, Feenstra, and Ware 1989; Gabaix 1999; Spector 1999; Bourguignon and Verdier 2000).

Institutions are part of the preceding framework precisely because they often define the set of feasible redistributions or the public decision rule in the absence of a consensus about the reform. As an example, I consider an extremely simplified representation of the role of democracy and social insurance. The extent of democracy is certainly behind the political economy or public choice mechanism leading to acceptance or rejection of the reform. Whether the decision is made by a ruling elite

or involves a broader set of actors clearly matters for the outcome of the public decision process. But the extent of democracy may also determine the capacity to redistribute the gains from the reform. If a commitment is needed from one of the two classes to redistribute at some time in the future, or not to cease redistributing, it may be more easily enforced in a democratic system where reputation matters.

Another relevant institution is, of course, social insurance or, more generally, all the mechanisms that generate some automatic redistribution across income levels or occupational statuses, for example, from employed to unemployed.

Suppose, for instance, that some linear, progressive tax and benefit system is in place before the reform considered above, that all incomes are taxed at the marginal rate t , and that the proceeds are distributed equally to all citizens. If the two classes have the same demographic weight, the net gains $-a$ and $a + b$ become $-a(1 - t) + b/2$ and $(a + b)(1 - t) + tb/2$. So, the social insurance institution redistributes the total benefits from the reform more evenly across the population. This may be sufficient to ensure that both classes benefit.

What might the role of democracy be in this simple example? Assume now that the basic parameter t of the social insurance distribution is decided by a majority vote. It is well known that in this case the marginal rate of taxation increases with the level of inequality in society. This clearly reinforces the distribution of the benefits of the reform throughout the population and the probability that both classes will benefit.

Democracy and social insurance essentially ensure that the distribution of income or well-being will never diverge too much from that in some reference situation. Writing such institutions into the constitution may resolve distributional conflicts once and for all, and society may more easily and more systematically seize opportunities for aggregate enrichment. If these opportunities occur randomly, a society with these kinds of institutions should fare better in the long run than a society without them. But these institutions have costs in terms of economic efficiency, and these costs may have to be balanced against the benefit of more systematically seizing good opportunities for reform.

This leads to my last point, on how to analyze decisions about undertaking changes relating to such institutions. This kind of institutional reform is much harder than policy reform, and I am not sure general answers to this question are possible. Consider the case of a democratic society deciding whether to introduce social insurance that would cushion the effect of negative shocks on the income of the poor. Identifying who would lose and who would gain from this social insurance scheme requires imagining all possible shocks and events that might trigger the scheme, determining how much everyone would lose or gain in each of these circumstances, and assigning some kind of probability to the occurrence of each event. This clearly is a formidable task. Another problem, of course, is who should decide whether to make the institutional change. Paradoxical situations arise when the institutional change is precisely about the mode of public decisionmaking.

As Rodrik mentions, big institutional reforms tend to take place in exceptional conditions. Historically, such conditions were social conflicts or threats of such con-

flicts. The continental European social insurance system, for example, was born partly out of Bismarck's fear of rising socialism. But we clearly cannot wait for such circumstances for desirable institutional changes in developing countries. What should be done? The development community must make a strong general case for such reforms, with due recognition of the specific conditions in a country when proposing any institutional change. But the task of evaluating these changes from different perspectives in a society has to be undertaken one way or another. Rodrik and Collier appear to recommend that the international and national development communities make this evaluation their top priority.

I am happy to concur with them. However, we probably shall not complete this quest for a long time, and until we do, policy reforms within existing and possibly unsatisfactory institutions will remain on our agenda. We will need to improve our capacity to deal with them.

References

- Bourguignon, François, and Thierry Verdier. 2000. "Is Financial Openness Bad for Education? A Political Economy Perspective on Development." *European Economic Review* 44: 891–903.
- Feenstra, Robert. 1987. "Incentive-Compatible Trade Policies." *Scandinavian Journal of Economics* 89 (3): 373–87.
- Gabaix, Xavier. 1999. "Technical Progress Leading to Social Regress and the Losses on Trade." Harvard University, Department of Economics, Cambridge, Mass.
- Lewis, Tracy, Robert Feenstra, and Roger Ware. 1989. "Eliminating Price Supports: A Political Perspective." *Journal of Public Economics* 40 (2): 159–85.
- Spector, David. 1999. "Is It Possible to Redistribute the Gains from Trade Using Income Taxation?" Massachusetts Institute of Technology, Department of Economics, Cambridge, Mass.

Comment on “Development Strategies for the 21st Century,” by Dani Rodrik

Fukunari Kimura

Japanese economists have claimed for some time that import substitution policies sometimes work well, that institutional diversity is important, that gradual policy reform is likely to be better than a “big bang,” and that applying a single policy prescription to all countries can be dangerous. Dani Rodrik lucidly presents views along the same lines, and I am sure that many scholars in Japan and elsewhere would happily acknowledge this endorsement of their claims by a prominent mainstream economist.

Rodrik argues that market incentives are critical to economic development, that these market incentives need to be underpinned by strong public institutions, that market economies are compatible with a diverse range of institutional arrangements, and that the greater the fit between market-oriented reforms and preexisting institutional capabilities, the higher the probability of success. He claims that the “augmented Washington consensus” takes only the first two into account. He proposes that we develop a new paradigm that also takes care of the latter two.

Rodrik’s article is intellectually provocative, and I agree with most of its claims about the relationship between markets and institutions. But I am concerned about throwing away the old paradigm before establishing a new one. Development is a very practical field. No matter how intellectually sincere, a notion that neglects its potential political economy implications will not necessarily bear good fruit in practice.

Take the Japanese experience with free trade. Trade theorists kept showing theoretical examples in which free trade was not optimal, even though we knew that free trade was “pretty good” in most cases. As a result, we were not very successful in convincing the public of the merits of free trade. This failure was due in part to our professional sincerity in exhibiting exceptional cases while not clearly presenting a rule of thumb. To guide policymakers, we need a benchmark model even if it does not always apply.

The issue discussed here has a different scope, of course, but the relationship between theory and practice is nevertheless similar. Because we always have to fight pressure from people who resist policy reform, we desperately need a simple, common, and

Fukunari Kimura is professor of economics at Keio University.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

seemingly objective guideline (rule of thumb). If we said, "Because institutions differ across countries, we must proceed case by case," we would provide excuses for avoiding policy reform. Such a statement might be intellectually honest but politically inadvisable in some parts of the world.

As Rodrik asserts, we have recently acquired a much better understanding of the complementarity between markets and the state, the importance of institutions, and the implications of institutional diversity. The question is how economists should deal with the institutional aspects of development. We now know that institutions are diversified for good reason and crucial to economic development. But we economists do not always know enough about institutions. So, what do we do? Should we discard all the old ways?

We do not have to lose all confidence in economics. Nor should we assume that only someday will we economists understand and be able to manipulate all development issues, including institutions. We must first do what we can do. Some issues are still straightforward, and our traditional paradigm can still provide a robust rule of thumb.

Rodrik acknowledges a strong confidence in the market. And there are still a number of basic policy issues that we can deal with using the traditional methods of economics. Although current economics cannot handle every detail of development policy, it can still provide general policy guidelines.

How can we learn from history? Rodrik claims that import substitution policies are not statistically associated with poor growth performance. But in future research we must digest the casual observation, record the possible economic causality, and specify when import substitution works well, when it does not, and when it does not matter. The traditional deductive approach can still be effective in drawing policy implications from history.

The same is true for international economic integration. The weak statistical association between economic integration and growth performance does not necessarily mean that economic integration is unimportant. There is still much room for analyzing the economic implications of open-door policies. We have learned a great deal from the Asian crisis, for example, about the implications of capital account liberalization.

As for Rodrik's market-sustaining public institutions, we have to reorganize issues by classifying them into what we can and cannot handle with traditional economics. Regulatory institutions can be handled to a great extent by traditional microeconomics. But it seems practical to deal with macroeconomic stabilization policies using an institutional approach, as Rodrik suggests, rather than rigorous microeconomic models. For social insurance institutions, setting the objective function is the key. It is still useful to separate efficiency issues from such social objectives as income distribution, as economists traditionally do to fight political pressure.

In sum, I agree that our improved understanding of the relationship between markets and institutions should encourage a substantial modification of our development strategies. But neither the traditional approach nor the augmented Washington consensus is useless or harmful, and neither should be immediately discarded. Many "simple" issues can still be handled using the traditional approach. And when we are reasonably confident, it is important to present a rule of thumb. We economists should neither underestimate nor overestimate our role in handling institutions. By considering the political economy implications of our work, we can still play a crucial role in economic development.

Rethinking Aid

Jan Willem Gunning

This article reviews the recent debate on the role of aid in providing finance, in transmitting knowledge, and in changing policies in recipient countries. It argues that where economic reforms are recent—as in much of Africa—donors can play an important signaling role. This potential remains untapped because aid allocations have largely been based on political considerations. Donors fear “aid dependency” and have therefore pushed for the substitution of domestic taxation for aid. The cost of taxation may instead justify the use of aid for tax relief, particularly where the cost of taxation is likely to decline with development.

For much of the past 50 years development aid was seen as simultaneously fulfilling two roles: first, financing projects in poor countries that lacked access to world capital markets and, second, raising the returns to such investment through the donor’s role in project selection and design. But since the early 1980s the role of aid has become less clear, partly because these two traditional roles have come to be questioned. The growing integration of many developing countries with global financial markets has made the provision of finance seem less relevant, while the fungibility critique has made donors’ efforts in project selection seem an exercise in self-delusion. The rationale for aid became further blurred when structural adjustment lending introduced a third role: the attempt to use aid to buy policy reform in developing countries.

The past few years have seen a radical critique of aid. This recent debate has addressed all three aspects of aid—its role in providing finance, transmitting knowledge, and changing policies in recipient countries. The most influential recent critique of aid is the World Bank’s *Assessing Aid* report (World Bank 1998). The argument in that report is based on two propositions. First, aid can be effective in

Jan Willem Gunning is professor of economics at the Free University, Amsterdam, and director of the Amsterdam Institute for International Development. The author is grateful to Chris Adam, Paul Collier, Jan Dehn, David Dollar, Chris Gilbert, and David Vines for helpful discussions.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

raising growth—but only in a good policy environment. Second, aid cannot buy such an environment.

There has also been a revival of interest in two other functions of aid: the provision of insurance through compensatory schemes and the use of aid as a commitment and signaling device. Where economic reforms are recent—as in much of Africa—risk assessments are slow to adjust, so for a time the private response to reform may be weak. In these circumstances aid can play an important signaling role. Because the allocation of aid has been dominated by political considerations, this potential remains unexploited.

This article reviews the debate on the roles and effectiveness of aid. First it considers the role of aid in providing finance. Current evidence suggests that aid is effective in raising economic growth only in good policy environments. The article then considers the effectiveness of aid in securing such an environment through ex ante conditionality. There is now overwhelming evidence that such conditionality does not work. The paper argues that donors should switch to ex post conditionality (selectivity). Donors are concerned about the possibility of “aid dependency.” They have therefore pushed for the substitution of domestic taxation for aid. I suggest that the cost of taxation may be atypically (and temporarily) high in very poor countries. This might justify the opposite substitution—that is, the use of aid for tax relief. Finally, the article considers the argument that the transfer of knowledge to developing countries should be tied to the provision of finance. It argues that there is no need for such “bundling” but that there is a strong case for clearly separating these two functions.

Aid as Finance

If the role of aid is to provide finance (rather than to change policies) then aid would be effective if it stimulated economic growth, taking policies as given. Much of the recent debate on aid has focused on the first of the two *Assessing Aid* propositions: that aid cannot be effective in a poor policy environment. This proposition is, of course, intuitively appealing. There is now empirical support from two types of evidence. First, micro evidence indicates that the returns to aid-supported projects are affected by macroeconomic policies. In the early 1980s accumulating anecdotal evidence of this nature was an important argument in favor of adjustment lending: it was thought that if project returns were low because of the policy environment then donors should focus on improving policies rather than on project design. There is now more systematic evidence—notably that of Isham and Kaufmann (1999), who analyzed World Bank-supported projects and found that macroeconomic policies had a significant effect on their returns.

The second type of evidence comes from growth regressions and relies on macroeconomic data. Burnside and Dollar (2000) included both aid and an aid-policy interaction term in a growth regression and found that aid alone was not significant but that the interaction term was. The Burnside and Dollar paper has attracted much attention since it first appeared as a working paper in 1997. The econometric basis

for the conclusion that aid works, but only in a good policy environment, has been attacked in a number of papers (for example, Hansen and Tarp 1999; Lensink and Morrissey 1999; and Lensink and White 1999).¹ But in the published version of their paper Burnside and Dollar use extensive sensitivity analysis to test the robustness of their result.

A particular interesting critique of the Burnside and Dollar thesis is that of Guillaumont and Chauvet (1999). Their hypothesis is that the effectiveness of aid (in terms of growth) depends not only on the quality of the policy environment but also on the country's "structural vulnerability." They measure vulnerability as a weighted average of the instability of agricultural value added (as a proxy for climate shocks), the instability of the income terms of trade, the trend value of the terms of trade, and initial population. Adding this vulnerability variable (and its interaction with aid) to the Burnside and Dollar (2000) growth regression, Guillaumont and Chauvet find that the aid-policy interaction term is no longer significant—while the aid-vulnerability interaction term is significant. On this basis they conclude that aid effectiveness does not increase with the quality of the recipient country's policies but rather with its vulnerability.

How should this result be interpreted? Consider first the trend in the terms of trade. In the absence of aid, a secular decline in the terms of trade might well induce an increase in savings: agents would accumulate assets in order to smooth consumption over time. (Note that at a later stage the effect on the savings rate would be negative: agents would maintain consumption levels by drawing down the assets they had accumulated initially.) However, if aid were available to compensate for the terms of trade loss then agents would have less incentive for such intertemporal substitution. To that extent the effect of aid on the savings rate (and hence on the growth rate) might well be negative, at least initially. The trend in the terms of trade component of the vulnerability therefore provides no clear explanation for the aid-vulnerability interaction term.

Now consider the two instability measures. Guillaumont and Chauvet (p. 8) argue "that aid has to allow countries to overcome their vulnerability, to face shocks in better conditions, or at least to (partially) compensate for their effects." This suggests that in shock-prone economies aid could fulfill two roles, addressing either the ex post or the ex ante effects of volatility on growth. Aid could provide insurance, offsetting the effects of negative shocks once they occurred (ex post), through compensatory finance. An example is the European Union's Stabex program. Alternatively, aid could assist countries ex ante, not in response to actual shocks but in response to an assessment that the country was facing uncertainty.

This distinction between the ex ante and ex post effects of volatility has long been prominent in theoretical work on trade shocks (see Bevan, Collier, and Gunning 1991 and Collier and Gunning 1999c) but only very recently has it begun to be tested. Dehn (2000) constructs an uncertainty measure as the GARCH (generalized autoregressive conditional heteroscedasticity) conditional variance of one-step-ahead forecast errors, regressing first differences in the price series on their own lag, a second lag in levels, and a time trend. Essentially this purges the series of both level

and differences information so that the residual can be considered as a measure of the uncertainty faced by an agent who uses all available information. Dehn defines trade shocks as the actual occurrence of a “large” forecast error (using 2.5 percent cutoff points). Adding both variables—the ex ante measure of uncertainty and the ex post measure of shocks—to a canonical growth regression, he finds (for a sample of 56 countries in 1970–93) that positive shocks do not have a significant growth effect, that large negative price changes lower the growth rate significantly, and, most interestingly, that uncertainty in itself does not affect growth. Hence what reduces growth is not vulnerability in the ex ante sense—the prospect of volatile world prices—but the actual realization of negative shocks.²

This does not invalidate the Guillaumont-Chauvet result, but it does suggest a refinement of their policy conclusion that the allocation of aid across countries should be based not only on the quality of policies but also on countries’ vulnerability. Recall that their vulnerability variable aggregates over quite different effects. Aid is not likely to be more effective in economies exposed to commodity price uncertainty, simply because such uncertainty in itself does not reduce growth. But negative shocks do matter for growth, and this suggests the need to reassess the insurance role of aid. The function of aid in shock-prone economies would be to relax borrowing constraints for countries that had experienced negative shocks.

Existing schemes such as Stabex provide compensatory finance to governments rather than to private producers. They thereby unintentionally reinforce (through Dutch disease effects) rather than dampen negative shocks for producers. When compensation for negative shocks is paid to governments, producers may lose twice: once directly through the reduction in producer prices and once indirectly because the government’s spending of the compensatory finance would drive up the relative prices of nontradables, thereby reinforcing the terms of trade loss suffered by the producers of the commodity affected by the shock. Aid would be more effective if it were directed at private producers, enabling them to buy put options. (Collier and others 1999 propose a redesign of Stabex along these lines.) But until recently donors have shied away from such a use of aid, believing that producers of commodities with volatile prices (such as coffee or sugar) should be encouraged to abandon these activities rather than supported through risk sharing.

With this qualification, aid can be effective when it takes the form of compensatory finance (ex post) for countries subject to negative trade shocks. In the more general case aid effectiveness appears to require a good policy environment. That is, of course, the rationale for the use of conditionality in adjustment lending—to effect policy reform in recipient countries.

Conditionality

In the early 1980s structural adjustment lending extended conditionality from projects to a wide range of government policies. This has led to theoretical work that calls into question the power of donors to change policies through conditionality, to

empirical research on the effectiveness of conditionality, and to proposals to reform conditionality. I consider these three developments in turn.

The Limited Power of Donors

The effectiveness of *ex ante* conditionality has long been questioned. One line of criticism in the theoretical literature stresses that if adjustment finance is temporary then conditionality can be effective in changing policies at best only temporarily. It may well be possible to use aid to “buy” policy reforms from unwilling client governments, but those reforms will not be sustained if aid is temporary.

For example, a government that prefers protection over free trade may well be induced by the offer of aid to liberalize its trade policy, but it would then have an incentive to reverse that reform once the aid runs out (see Rodrik 1989; Collier and Gunning 1992; and Killick 1998). The government’s announcement that its trade reform was permanent would be recognized as “time inconsistent” and hence be considered incredible. This would explain both why policy changes undertaken at the insistence of donors have often been short-lived and why the response of private agents to the announcement of adjustment policies has often been weak. To explain why policy reversals frequently occur even in repeated donor-recipient interactions, it is usually argued that a donor threat to punish policy reversals by denying future access to adjustment lending is not credible given the incentives for donor staff to continue lending (see, for example, Killick 1998). Arguments for *ex post* conditionality are often inspired by models in which time inconsistency generates policy reversals.

Time inconsistency is, of course, not the only reason for the failure of conditionality. A good example from the political economy literature is provided by Adam and O’Connell (1999). In their two-period model private agents make investment decisions at the beginning of the first period, after the government has announced its tax and transfer policies.³ Agents can invest in two activities, one of which is taxed. The government collects the revenue from this (distortionary) tax, receives aid from donors, pays for exogenous government expenditure, and uses any remaining revenue for transfers to a favored group. It chooses the tax rate and the amount of transfers so as to maximize the welfare of this group.

The size of this favored group is critical. If it represents a small portion of the population, then it is optimal for the government to engage in transfer payments. If, however, the government is “sufficiently representative” (in the sense that the favored group exceeds a critical portion of the population), then it will be benign. Despite being interested only in the welfare of its clients, the government will make no transfers to them and will set the tax rate at the minimum needed to finance the part of exogenous spending not financed by aid.⁴

The effect of (unconditional) aid differs radically in these two types of economies. In the first case (where the government has been captured by a small rent-seeking elite) any increase in aid will be used entirely for transfer payments. In the second case a (small) increase in aid will not induce any transfers: transfers will remain constant at zero and aid will be used exclusively for tax relief. To the extent that poverty is con-

centrated in countries with narrowly based governments, this model illustrates what is sometimes described as the problem of selectivity: aid (to the government) is ineffective in the very economies where the depth of poverty makes it seem most desirable.

This simple model shows both how aid affects growth and why aid effectiveness depends on the policy regime. In the economy with a sufficiently representative government, the quality of the policy environment is high in the specific sense that the government chooses to make minimal use (given the level of expenditure) of distortionary taxation. In this economy aid raises growth by inducing a reduction in the tax rate and thereby an increase in investment in the high-return (taxed) activity. Conversely, in the nonrepresentative (or more accurately, not sufficiently representative) economy aid affects growth only by raising transfers to the favored group, which then increases its saving to smooth consumption between the two periods.

When aid is unconditional, the recipient government is, of course, free to choose the tax rate that is optimal given its objective of maximizing the welfare of the favored group. Under conditionality, however, the tax rate and the amount of aid are simultaneously determined in bargaining between the donor and the recipient where the donor will be concerned by the effect of distortionary taxation, through its effect on investment, on consumption in the second period. If the government is not sufficiently representative, there are “gains from aid”: conditionality can make both parties better off. (For the recipient government this is so because without conditionality it would have received no aid at all since the donor would have realized that aid would have been wasted on transfers, without reducing tax rates.) Adam and O’Connell show that whether conditionality is needed to achieve gains from aid or, more narrowly, Pareto efficiency, depends on the size of the government sector (as measured by the amount of exogenous government expenditure) and on the size of the group favored by transfers.

This second version of the model (in which aid is conditional) illustrates the function of conditionality but not its possible failure through policy reversals. While the model distinguishes two periods it is essentially static: policy decisions are made once and for all at the beginning of the first period. Hence a problem of time inconsistency cannot arise.

Evidence on the Effectiveness of Conditionality

Policy reversals are, however, a central concern in the empirical literature. Recall that the second proposition in the *Assessing Aid* report was that aid cannot buy policy reform. Empirical evidence on this proposition is largely in the form of case studies, although there have been many econometric attempts, notably that of Burnside and Dollar (2000). In assessing this evidence it should be noted that for conditionality to be effective, it is necessary but not sufficient that the recipient government change its policies in a way deemed desirable by the donor.

This may be seen by noting that when a donor tries to buy policy reforms through conditionality there are four conceivable outcomes (Gunning 2000):

- The reforms are not adopted, so conditionality has obviously failed.
- The reforms are adopted but this would also have happened without aid.
- The reforms are adopted and would not have been adopted in the counterfactual case (if no aid had been offered), but the effect is only temporary: the reforms are ultimately reversed (the outcome predicted by time inconsistency models).
- The reforms are adopted and sustained, which would not have occurred without aid.

Clearly, conditionality is truly effective only in the last case. It achieves nothing in the first and the second case (although only in the first case is this obviously so) while in the third case it is only temporarily effective. Regression analysis runs into an obvious limitation here: it can establish whether aid flows are accompanied by policy reforms, and it may be able to establish whether reforms are reversed. But it cannot convincingly distinguish between situations where aid was instrumental or only incidental to policy reforms. Regression evidence of a link between aid and policy quality at best shows that a necessary condition for effective conditionality is satisfied. Case studies go some way toward overcoming this problem by using in-depth, country-specific information to construct a credible counterfactual. However, here too there is a methodological problem: donors have an incentive to exaggerate their influence over policy changes while the government has the opposite bias. Hence case studies may be biased in either direction, depending on which policy accounts they rely on.

Most econometric studies in this area focus on whether donor finance is effective at changing policies in recipient countries. Dollar and Svensson (2000) focus instead on the effects of the donor's preparation and supervision activities. Using data from the World Bank's Operations Evaluation Department (OED), they attempt to explain a program's success or failure (as assessed *ex post* by OED) by the donor's efforts. There turns out to be no such relationship: the authors "find no evidence that any of the variables under the World Bank's control affect the probability of success of an adjustment loan" (p. 4).⁵ These variables are measures of staff input, not of the aid provided. The authors, as they point out, therefore do not test the proposition that aid can buy policy change. What they do test is whether the donor's activities that complement finance—the design of a set of conditions, the technical assistance provided, and the staff efforts to twist the government's arm—determine success or failure. While the evidence is only for the World Bank, it seems unlikely that other donors would have been more effective.

The case study evidence is growing rapidly; see, for example, the country studies in the external evaluation of the International Monetary Fund's (IMF's) Enhanced Structural Adjustment Facility (Botchwey and others 1998), the studies in Killick (1998), the World Bank study of aid in Africa (Devarajan, Dollar, and Holmgren 2000), and the recent evaluation of Swedish program aid (White 1999). All this evidence strongly suggests that conditionality is not effective. For example, the Devarajan, Dollar and Holmgren study of aid and policy reform in 15 African countries finds that while all of them got large amounts of program aid, only 3 (Mali, Ghana and Uganda) reformed successfully.

Many governments simply do not adopt the reforms favored by donors (the first possible outcome suggested above). For example, in the final phase of the Kaunda regime in Zambia, the government reached an agreement with the IMF on a program “designed to create a diversified and market-oriented economy” (internal IMF document, as cited in Botchwey and others 1998, p. 95). The government had no intention of creating such an economy. Indeed, price controls were formally abolished but continued to be enforced, and efforts to accord a larger role to the private sector were resisted effectively (Botchwey and others 1998). More recently again in Zambia, the government long resisted donor pressure to privatize copper mines. There are also many examples of aid, far from inducing reforms, actually postponing them. (Devarajan, Dollar, and Holmgren 2000 document this for the Democratic Republic of Congo, Kenya, Nigeria, and Tanzania.)

There is also evidence of reforms being adopted but not because of donor pressure (the second possible outcome suggested above). In Vietnam donors made an important contribution by informing the policy debate but were apparently virtually powerless to force through reforms against the wishes of the authorities. As van Donge and White (1999, p. 33) stress: “The pace and direction of reform is determined by Vietnamese politics.”

The literature also documents many cases where reforms were “owned” by the government, so donor pressure was not essential. At one extreme reforms were even undertaken without donor involvement, as was the case in the early phases of liberalization in Burkina Faso, Eritrea, Mozambique, Uganda, and Vietnam. In Uganda key reforms (abolishing price controls, liberalizing the foreign exchange market, privatization) were all undertaken on government initiative. While donors, understandably, like to take credit for the success of reforms in Uganda, there have been very few cases of substantial disagreement on policy issues between donors and the government (Ddumba-Ssentamu and Dijkstra 1999, pp. 91–92). Kasekende and Atingi-Ego (1999) argue convincingly that Uganda’s successes cannot be attributed to donor pressure.

This is not to say that the involvement of donors in these countries was not important or helpful; it often was. But the contribution of donors was to help build the case for reform through policy dialogue rather than to buy reform with aid. In addition, donor support eased the implementation of reform programs (Devarajan, Dollar, and Holmgren 2000).

The third possible outcome, involving policy reversals, has been especially common in Africa. “Program interruptions” (in most cases involving policy reversals) were the main concern in the IMF’s internal evaluation of its Extended Structural Adjustment Facility (IMF 1997). A well-known case is the liberalization of maize marketing in Kenya, a reform that was repeatedly undertaken, with each attempt ending in a policy reversal. Oyejide, Ndulu, and Gunning (1999) document that trade liberalizations have been reversed in 7 of 10 African countries, in many cases (including Kenya) more than once.

In the fourth possible outcome conditionality is effective in the sense that donor pressure was essential and the reform was sustained. This case seems fairly rare. In an exhaustive evaluation of Swedish program aid one of the few examples

is the liberalization of Mozambique's cashew market—a reform that donors effected despite strong government objections (White 1999; Gunning 2000).

Responses to the Evidence on the Failure of Conditionality

The central idea of structural adjustment lending is that aid can buy policy reform. Having committed huge resources to this idea for two decades, one would expect donors to be shocked by the accumulating evidence that conditionality has failed. Instead that message seems to be accepted with remarkably little dissent. The extent of the recent change in thinking on aid can be measured by looking at the rationale Rodrik (1996) gave five years ago for (multilateral) lending. Rodrik saw the exercise of conditionality as a key rationale for aid. Donor confidence in this role for aid appears to have evaporated in a very short period.

If donors accept the evidence that *ex ante* conditionality does not work, they can either redesign the aid contract to improve it or switch to *ex post* conditionality. The *ex ante* approach weakens ownership, the *ex post* approach reinforces it. Much of the theoretical literature is concerned with establishing the feasibility of an effective, incentive-compatible aid contract. But the problem with *ex ante* conditionality goes beyond feasibility: one may well question its desirability. If donors are to succeed in achieving objectives that are not (fully) shared by the government,⁶ they inevitably will undermine the government's accountability. In the limit conditionality is effective in the same way as colonialism and hence suffers from the same objection. Stiglitz (1998, pp. 10–11) eloquently describes the debilitating effect of *ex ante* conditionality:

Rather than learning how to reason and develop analytical capabilities, the process of imposing conditionality undermines both the incentives to acquire those capacities and confidence in the ability to use them. Rather than involving larger segments of the population in a process of discussing change—thereby changing their ways of thinking—it reinforces traditional hierarchical relationships. Rather than empowering those who could serve as catalysts for change within these societies, it demonstrates their impotence. Rather than promoting the kind of open dialogue that is central to democracy, it argues at best that such dialogue is unnecessary, at worst that it is counterproductive.

In summary, *ex ante* conditionality has proven ineffective. It may well be feasible to make it more effective; however, this would reinforce the undesirable effects stressed by Stiglitz. The alternative is *ex post* conditionality.

Selectivity

If aid effectiveness depends crucially on the quality of the policy environment and donors are powerless to change policies with aid, then all they can do is improve efficiency through selectivity in the allocation of aid. Donors would treat policy

regimes as exogenous and would bias the allocation of aid in favor of countries with good policy environments.

The *Assessing Aid* report does not shy away from this logical implication. Some bilateral donors, notably the Dutch, are rapidly moving in this direction. The World Bank itself seems much more reticent. As so often happens in debates on development, there is a wide gap between the world of ideas and reality.

It should be noted that while selectivity can be defended even if policy remains exogenous, its beneficial effect may be reinforced through an incentive effect. This would be the case if governments recognized a donor's new allocation rule and adopted economic reforms partly on the expectation that they would be rewarded with aid (Collier and others 1997).⁷ The policy would then become endogenous.

Alesina and Dollar (2000) show that the allocation of bilateral aid is driven by political and strategic considerations, such as colonial ties and United Nations (UN) voting records. By contrast, recipients' economic policies and political institutions are relatively unimportant determinants of aid flows. The authors find, for example, that "a non-democratic former colony gets about twice as much aid as a democratic non-colony" (p. 21).

The importance of noneconomic objectives is quite striking. For example, the Alesina-Dollar results indicate that donors reward openness: open economies (on the Sachs and Warner 1995 index) get twice as much aid as closed ones. But this effect is swamped by variables measuring colonial ties and UN voting records (which are not correlated with openness). When allocations are based on the policies adopted by recipient governments, the relationship is often perverse: aid tapers off in successful economies (Botchwey and others 1998; Collier and Gunning 1999b; Devarajan, Dollar, and Holmgren 2000). This appears to be based on donors' mistaken perception that the return to aid in such economies is low because they already attract private investment.

Selectivity can be seen as the most radical form of ex post conditionality because it treats the policy regime as exogenous and hence abandons all pretense that aid can buy policy change.⁸ The selectivity proposal has encountered (at least) three objections.⁹

The first is that selectivity will leave poor people living under governments with bad policies to fend for themselves—that is, without aid. This objection is hard to understand: aid to such governments is unlikely to benefit the poor anyway, and it may well harm them by financing bad policies. (Recall that Devarajan, Dollar, and Holmgren 2000 find that aid often maintains bad policies.) Conversely, selectivity does not rule out that aid is channeled directly to poor people rather than to their governments.

The second objection is that countries with good policies do not need aid. This is simply false. The evidence on risk ratings indicates that reputations die quite slowly. Hence domestic savings may remain low for some time in good policy environments, reflecting uncertainty as to whether the new policy regime will be sustained. There may therefore be a phase in which a country has already undertaken wide-ranging reforms but has not yet attracted substantial private capital. Also, in poor countries the cost of taxation is typically quite high. Encouraging countries

with good policies to substitute domestic taxation for aid (a favorite proposal of critics of “aid dependency”) may therefore be quite damaging.¹⁰ I return to this point in the next section.

The most serious objection to selectivity in aid allocations is that it may encounter the same problems as *ex ante* conditionality—namely, if it leads to donor-recipient bargaining over the quality of the policy environment. Selectivity ties a country’s aid allocation to an assessment of its policies. Some aspects of a “good policy environment” can be defined unambiguously in objective terms, but in many cases an element of judgment is inevitable. The donor’s judgment may, of course, be challenged by the recipient government. In that case the government and the donor would return to the sort of bargaining that now characterizes *ex ante* conditionality. Selectivity would then conflict with ownership in the sense that donors would try to impose their judgment on the relative importance of various aspects of the policy regime. In the limit there would simply be no scope for ownership: aid would be allocated to countries that had adopted a list of donor-favored policies. This objection is the main reason for favoring outcomes over policies as the basis for *ex post* conditionality.

There is an active debate on whether conditionality should be based on outcomes or on policies adopted by the recipient government, as has been the rule in structural adjustment programs. Practical people tend to favor policy-based conditionality, largely because of the long lags between policy decisions and their results—say, in terms of poverty alleviation. Incentive effects would be weak if governments could not expect to be rewarded (or punished) quickly for their actions. Rewards might actually be reaped only by a successor government.

The theoretical literature puts more emphasis on outcomes. Azam and Laffont (1998) develop a principal-agent model in which the principals (the rich in the North) try to induce the rich in the South to raise poor people’s consumption, which the donors value altruistically. In their model an optimal contract may make aid conditional both on policies and on outcomes, the latter because of incentive effects.¹¹

An obvious objection to outcomes-based conditionality is that government has imperfect control over outcomes. In particular, outcomes might reflect exogenous shocks—which might lead a passive government to receive less aid simply because the country experienced a negative shock. Collier and others (1997) suggest correcting for this by regression analysis. They introduce in a growth regression a number of determinants outside government control—such as terms of trade shocks and landlockedness. Using such a regression, an aid allocation based on growth would correct observed growth rates for the estimated effect of the nonpolicy determinants. For example, a country that experienced a negative trade shock in the period under review would have its growth rate adjusted upward by adding to the observed rate the shortfall resulting from the shock, using the estimated coefficient.

This correction may suffer from omitted variable bias if policies are not included and if policies are correlated with some other growth determinants.¹² If the regression does include policy variables then the adjusted growth rate (on which the aid allocation would be based) would be equal to the part of the growth rate not

explained by the nonpolicy variables—that is, it would equal the error term plus a weighted sum of the policy variables with the estimated regression coefficients as weights. In this sense the difference between policy and outcomes-based aid allocations would be illusory. In the outcomes-based approach, aid would actually be allocated on the basis of the policies the government had adopted.

There are, however, two important differences, one economic and the other political. The economic point is that the policy-based approach rewards policy choices on the basis of their average effect in the sample used for the regression. The outcomes-based approach would include the error term in the measure of success. Whenever a government had been able to realize higher returns to particular policies than the sample average, that would be picked up in the error term and hence be rewarded. The outcomes-based approach would therefore encourage the government to make its own decisions about the relative importance of reforms, giving, say, more weight to trade reform and less to fiscal rectitude if that seemed appropriate in the specific circumstances of that country. The political point is that an outcomes-based allocation promotes accountability by signaling that donors have no involvement in policy choices, being interested only in results.

Finally, under selectivity aid could serve both a signaling and a restraint role. Signaling would occur once it was recognized that aid allocations were tied to success. Private agents would then be able to economize on monitoring by observing aid allocations instead. Similarly, the government could use donors as an agency of restraint: if the aid allocation rule was credible, the government could resist pressures for policy reversals by pointing out that such reversals would lead to a loss of aid. This would help lock in policies.

Aid or Taxation?

Research on aid effectiveness has brought into sharp focus the relationship between aid and taxation in recipient countries. Where donors have little control over the composition of government spending (because of the fungibility of project lending and the limited effectiveness of conditionality in program aid), aid increases government spending and reduces taxation.¹³

That a recipient government responds to aid with such a combination is, of course, rational. Domestic taxation has both political and economic costs (in the absence of lump-sum taxation). It is therefore optimal for the recipient government to increase government expenditure and to reduce taxation until in the new equilibrium with aid the marginal costs of taxation is again equated (but now at a lower level) to the marginal benefits of government expenditure (see, for example, Collier 1999). Advocates of the “aid dependency” position argue, however, that increases in government expenditure should be accompanied by increased taxation and reduced aid. There is no rationale for this view in static donor-client models as long as the marginal benefits of government expenditure are decreasing and the marginal costs of taxation are increasing.

The real issue appears to be dynamic: how will the cost of taxation change in the process of development, and what does that imply for aid policies? Taxation involves two types of costs: the cost of collection (the cost of the administrative machinery needed to generate revenue) and the welfare loss imposed on the economy if lump-sum taxation is not feasible so that taxation reduces private agents' income, not only directly but also indirectly by giving them incentives to change their decisions.

First consider the welfare cost of distortionary taxation. This is likely to fall over time, as the country develops. This may be illustrated with a very simple two-period model. Private agents inherit a capital stock k and decide in the first period, after the government has announced the tax rate, t , how much of their first-period income $f(k)$ to save (i). Agents will choose the optimal level of investment so as to maximize the present value of utility in the two periods:

$$W = u(c_1) + u(c_2) \beta ,$$

where β is the discount factor, $c_1 = f(k) - i$ is consumption in the first period, and $c_2 = (1-t)f(k+i)$ in the second period.¹⁴ The first-order condition is:

$$u'(c_1) = \beta u'(c_2)(1-t)f'(k+i) ,$$

where primes denote derivatives. This condition simply equates the opportunity cost of investment (the utility of the consumption forgone in the current period) to its marginal benefit (the discounted utility of the increase in future consumption as a result of investment, taking into account that future output will be taxed). This condition implicitly defines the optimal level of investment as a function of the tax rate and the initial level of the capital stock. I assume that increases in the tax rate reduce investment.¹⁵ The effect of a change in the tax rate on the welfare (W) of private agents is given by:

$$\frac{dW}{dt} = -\beta u'(c_2^*)f'(k+i^*)$$

because the tax base is $f(k+i)$. Taking into account that the direct effect of an increase in the tax rate on tax revenue $T = tf(k+i)$ will be partly offset by the induced fall in investment:¹⁶

$$\frac{dT}{dt} f'(k+i^*) + tf'(k+i^*) \frac{di^*}{ft} .$$

Hence the effect on welfare of an increase in tax revenue is given by:

$$\frac{dW}{dt} = -\beta u'(c_2^*) \left[1 + t \frac{di^*}{ft} f'(k+i^*) \right] f'(k+i^*) ,$$

which indicates that increased taxation reduces welfare.

How will this cost of taxation change if the economy develops? A more developed economy will enter the first period with a larger capital stock. This will be

reflected in a higher level of investment and higher consumption in both periods. The increase in c_2^* will, of course, reduce the marginal utility of consumption. This lowers the cost of taxation: the utility loss due to a given reduction in consumption as a result of taxation is lower at higher levels of income. However, the denominator (which measures the elasticity of tax revenue with respect to the tax rate) will also be affected. In this one-sector model the elasticity will (for iso-elastic utility and production functions) decrease, and this would tend to increase the cost of taxation. But for plausible parameter values the first effect (the declining marginal utility of consumption) dominates, so that the cost of taxation falls as the economy grows. Also, in reality economic growth will involve the development of new activities, so that increased revenue is not only achieved by raising tax rates but also by developing new tax handles. This increases the elasticity of tax revenue with respect to tax rates, thereby reinforcing the fall in the cost of taxation.

Hence growth alone may well reduce the cost of taxation, even without any changes in the ways taxes are collected: there is no need to appeal to a fall in collection costs.

In static models of aid and taxation the rationale for aid is that it raises public expenditure and reduces the cost of taxation. The dynamic argument that I have sketched reinforces this: if the cost of taxation falls with development then this justifies concentrating aid on countries in the early phases of growth. Ironically, donors have in recent years tended to move in the opposite direction, insisting on tapering off aid to recently stabilized economies (Botchwey and others 1998; Collier and Gunning 1999b). This donor obsession with domestic resource mobilization (substituting taxation for aid) may well be misguided.

Now consider the second type of cost of taxation, that associated with tax collection itself rather than with the induced distortion. It might be argued that the collection cost will fall over time as the tax authorities gain experience. Azam, Devarajan, and O'Connell (1999) model this as a learning-by-doing process by assuming that the marginal cost of tax collection is a decreasing function of the revenue collected in the past. A government that chooses to increase taxes will thereby give its tax inspectors the opportunity to learn to become more efficient, thereby reducing future collection costs. Clearly, donors would have to take such a learning process into account when deciding on their aid programs. If they offered aid today, they might undermine the learning process by inducing a reduction in taxation effort. As Azam, Devarajan, and O'Connell show, this might lead to an equilibrium outcome with high aid and low tax collection efficiency. Aid would have the usual short-run benefits, but it would undermine institutional development by depriving the government of cost-reducing tax collection experience.

Whether the main effect of high taxes is indeed such a beneficial learning process is doubtful. While learning-by-doing can obviously play an important role in lowering tax collection costs, it is not clear why higher tax rates would stimulate this process. What agents would learn might not be efficient tax collection but efficient bribery and tax evasion.

These issues must be resolved empirically. Sadly, there is very little evidence on the cost of taxation in poor countries.

A Knowledge Bank?

Donors, notably the World Bank, have often argued that aid is an important vehicle for the transfer of knowledge, both in project and program design—that is, that knowledge is best transferred by tying it to aid or, in the jargon, by “bundling” finance and expertise (Gilbert, Powell, and Vines 1999). Bundling is necessary, the argument goes, because recipient governments will not accept advice if it is not accompanied by finance. There are three objections to this position.¹⁷

First, it is not supported by the evidence. In structural adjustment lending advice is, of course, bundled with finance. I have already noted the Dollar and Svensson (2000) finding that the use of Bank resources—in terms of staff input for the preparation and supervision of adjustment loans—has no effect on the success or failure of adjustment programs. While knowledge transfer is not the same as staff efforts on preparation and supervision, the two should be highly correlated. Hence, whatever bundling does, it cannot be said to ensure that knowledge is transferred in a form that contributes to program success.

Second, bundling is equivalent to the tying of aid and has the same disadvantage: the recipient government cannot choose between suppliers. Donors differ, of course, in the expertise on which they can draw when giving technical advice (if knowledge was truly a public good then no need would be perceived to bundle finance and expertise) and there is no guarantee that any particular donor knows best. More important, tying undermines the recipient government’s accountability. When a government is not encouraged to develop its own position on the pros and cons of, say, a particular method of privatization but relies instead on force-fed donor advice, it can more easily deny responsibility for a subsequent failure.

Third, if donors accept the case for selectivity, it would be inconsistent to maintain bundling. The case for selectivity is that finance does not succeed in changing policies. There is no evidence that bundling improves this record. But there is evidence that knowledge transfer alone can build the case for reform. Indeed, the highest payoff to technical advice on program design appears to come in the pre-lending phase (World Bank 1998; White 1999). Devarajan, Dollar, and Holmgren (2000) find that in two of the three successful reformers in their African sample, Ghana and Uganda, the role of aid was largely to provide ideas in the pre-reform stage, rather than to provide finance. Indeed, both countries appear atypical for having received little financial aid in their pre-reform phases.¹⁸

Unbundling seems particularly pertinent for program lending. If ownership is to be encouraged by letting governments free in designing their programs then evidently they should be left free in deciding what information they require and from whom that information is to be obtained. Governments would then hire technical expertise themselves, either from consultancy firms or from donor agencies. Clearly, this would radically change current donor practices.

Donors can transmit knowledge not only to recipient governments but also to private agents, including foreign investors. Here the case for a donor role is much stronger. There is now a class of poor countries where, as a result of recent economic reforms, the policy environment can be considered adequate. However, since the reforms are recent, the private financial sector still has outdated information on these economies. Because many of these economies are small, particularly in Africa, the private sector has little incentive to incur the high costs of acquiring the information needed for accurate risk assessments. As a result many newly reformed poor countries are rated as more risky than is justified on the basis of fundamentals. They therefore fail to attract investment despite the far-reaching reforms they have adopted in the past decade.¹⁹ This is important because these newly reformed countries are critically short of private investment. In this context the rationale for a donor role is the underinvestment by private agents in information about the government's type (see Rodrik 1996).²⁰

In this view the "knowledge bank" would focus not on transmitting knowledge to but knowledge about developing countries. As before, the question rises whether this calls for bundling. Some have argued that signaling by a donor agency is credible only when the agency has some of its own money at stake.²¹ I am not convinced by this argument. In the case of bundling, investors might perceive the signal as noisy if they suspected that the donor's lending to the country would bias its reporting. It would then be difficult to distinguish between genuine positive news based on the donor's superior information and positive news that reflected the donor's wishful thinking. Far from bundling the two functions, one would therefore want to separate them quite clearly. This might provide a basis for a division of labor between donor agencies—for example, between the World Bank lending for long-run growth in stabilized economies and the IMF limiting itself to signaling.

But it should be noted that the noisiness of the signal provided by donor involvement in an economy is largely the result of current aid allocation practices. When these change in the direction of selectivity, the problem will be solved automatically: donor involvement will be limited to good policy environments and so will in itself be informative to private agents.

Conclusion

The recent debate on aid effectiveness has focused on three functions of aid: providing finance, transmitting knowledge, and changing policies in recipient countries. In reviewing this debate, I have stressed the ineffectiveness of aid in the third of these roles and the desirability of separating the other two roles more clearly.

Whether aid is effective in raising growth rates through the provision of finance has come to be questioned. Evidence from growth regressions suggests that aid is effective—but only in good policy environments. The econometric dispute surrounding this finding has not yet been resolved. However, the importance of the policy environment also emerges from microeconomic evidence on the returns to projects. Evidence on the effect of negative trade shocks on growth suggests an insurance role for aid.

There is now overwhelming evidence that aid is not effective in bring about policy reform. I have argued that rather than redesigning the aid contract to make ex ante conditionality more effective, donors should switch to ex post conditionality (selectivity). Under selectivity the allocation of aid is tied to success. This would enable donors to play an important signaling role in transmitting information on government policies and outcomes to private agents. Because aid allocations have been dominated by political considerations, the potential of such a signaling role has remain largely untapped. Where economic reforms are recent—as in much of Africa—signaling can play an important role.

I have suggested that the cost of taxation is likely to be atypically high in poor economies. Donor efforts to promote tax efforts in these economies may be misguided. The high cost of taxation and the likelihood that it will decline with growth may justify the use of aid for tax relief.

I have also argued that the transfer of knowledge to developing countries need not be bundled with the provision of aid. Indeed, the two functions should be separated if the case for selectivity is accepted.

Notes

1. Hansen and Tarp (1999) show that, controlling for investment and human capital, aid is not effective in raising growth. This test seems unreasonably severe: one would expect aid to affect growth partly, perhaps largely, through those two variables.

2. The finding that by itself uncertainty does not affect growth appears quite robust. Dehn (2000) shows that it holds for nine definitions of uncertainty and across different periods.

3. Adam and O'Connell (1999) abstract from credibility problems. In their model private agents do not doubt whether the government will maintain its policies in the second period—that is, once private agents have committed themselves through irreversible investment decisions. There is therefore no issue of time inconsistency in this model.

4. The intuition for this difference is that in the second case members of the favoured group lose more as taxpayers than they gain in transfers since transfers have to be shared by a relatively large number of people.

5. There is, however, evidence that causality runs in the opposite direction. The evidence suggests that once a program appears likely to fail, the Bank allocates more resources to supervision in an attempt to salvage it.

6. If the two parties are in full agreement on objectives, there is no need for conditionality.

7. An anonymous referee of Burnside and Dollar (2000) points out that this is similar to the Lucas critique: if the aid allocation rule were to change, then past evidence on the effect of aid on policies would no longer be relevant. I am indebted to David Dollar for this point.

8. Recall that when the aid allocation rule is clear, there may well be an incentive effect so that aid does promote policy change. However, selectivity does not rely on this effect.

9. This section draws on Gunning (2000).

10. See Collier and Gunning (1999b). Collier (1999) presents an excellent critique of the concept of “aid dependency.”

11. The model is static, so the issue of the lag between government actions and outcomes-based aid does not arise.

12. Guillaumont and Chauvet (1999) consider a particular case—namely, if policies are endogenous. They advocate correcting the outcome measure not only for the direct effect of,

say, a negative shock on the growth rate, but also for its indirect effect, through an induced policy change. This proposal treats observed policy responses as inevitable, as if the government had no choice. I prefer to err in the opposite direction.

13. While this is plausible, it need not be true. For example, recall that if the government is insufficiently representative in the Adam-O'Connell model, aid has no effect on taxation: it is used entirely for transfers to the government's supporters. However, this case is extreme. In general aid will at least in part be used for tax relief.

14. The functions u and f are assumed to be strictly concave.

15. For a utility function with a constant degree of relative risk aversion, R , this will be the case for $R < 1$.

16. Assume that the net effect is positive—that is, that the economy is on the increasing part of the Laffer curve.

17. Note that the first and third objection are specific to program lending; the second objection applies to both program and project lending.

18. This provides support for the position argued in *Assessing Aid* that aid to countries not yet committed to reform should take the form of ideas rather than money. The third country, Mali, did receive considerable aid in the pre-reform stage, but this appears to have played little role in the adoption of reforms. Côte d'Ivoire, now often classified as a successful reformer, had a long period of half-hearted reforms (Botchwey and others 1998).

19. See Haque, Nelson, and Mathieson (2000) on the determinants of risk ratings, Jaspersen, Aylward, and Kox (2000) on the relationship between risk ratings and private investment, and Collier and Gunning (1999b) on the case for a donor role in signaling.

20. The case for the IMF to adopt such a signaling role in poststabilization economies is argued in Botchwey and others (1998) and in Collier and Gunning (1999b).

21. This was the position of the IMF's board in response to the recommendation by an external evaluation (Botchwey and others 1998) that the IMF adopt a signaling role (based on limited monitoring) in poststabilization economies.

References

- Adam, Christopher S., and Stephen A. O'Connell. 1999. "Aid, Taxation and Development in Sub-Saharan Africa." *Economics and Politics* 11 (3) : 225–53.
- Alesina, Alberto, and David Dollar. 2000. "Who Gives Foreign Aid to Whom and Why?" *Journal of Economic Growth* 5 (1): 33–63.
- Azam, Jean-Paul, and Jean-Jacques Laffont. 1998. "Aid, Conditionality and Poverty Monitoring." Université des Sciences Sociales, IDEI and ARQUADE, Toulouse.
- Azam, Jean-Paul, Shantayanan Devarajan, and Stephen A. O'Connell. 1999. "Aid Dependence Reconsidered." Working Paper 99-5. University of Oxford, Centre for the Study of African Economies.
- Bevan, David, Paul Collier, and Jan Willem Gunning. 1991. "The Macroeconomics of External Shocks." In V. N. Balasubramanyam and Sanjaya Lall, eds., *Current Issues in Development Economics*. London: Macmillan.
- Botchwey, Kwesi, Paul Collier, Jan Willem Gunning, and Koichi Hamada. 1998. *Report by a Group of Independent Persons: External Evaluation of the ESAF*. Washington, D.C.: International Monetary Fund.
- Burnside, Craig, and David Dollar. 2000. "Aid, Policies and Growth." Policy Research Working Paper 1777. *American Economic Review* 90 (4): 847–68.
- Collier, Paul. 1999. "Aid Dependency': A Critique." *Journal of African Economies* 8 (4): 528–45.

- Collier, Paul, and Jan Willem Gunning. 1992. "Aid and Exchange Rate Adjustment in African Trade Liberalisations." *Economic Journal* 102 (July): 925–39.
- . 1999a. "Explaining African Economic Performance." *Journal of Economic Literature* 37 (March): 64–111.
- . 1999b. "The IMF's Role in Structural Adjustment." *Economic Journal* 109 (459): 634–51.
- . 1999c. *Trade Shocks in Developing Countries*. Oxford: Oxford University Press.
- Collier, Paul, Patrick Guillaumont, Sylviane Guillaumont, and Jan Willem Gunning. 1997. "Redesigning Conditionality." *World Development* 25 (September): 1399–1407.
- . 1999. "Reforming Stabex." *World Economy* 22 (5): 669–82.
- Ddumba-Ssentamu, John, and Geske Dijkstra. 1999. "Swedish Programme Aid to Uganda: An Evaluation." SIDA Evaluation 1999/17:6. Swedish International Development Authority, Stockholm.
- Dehn, Jan. 2000. "Commodity Price Uncertainty and Shocks: Implications for Investment and Growth." D. Phil. diss. University of Oxford, Centre for the Study of African Economies.
- Devarajan, Shantayanan, David Dollar, and Torgny Holmgren, eds. 2000. *Aid and Reform in Africa*. Washington, D.C.: World Bank.
- Dollar, David, and Jakob Svensson. 2000. "What Explains the Success or Failure of Structural Adjustment Programmes?" *Economic Journal* 110 (466): 894–917.
- Gilbert, Christopher, Andrew Powell, and David Vines. 1999. "Positioning the World Bank." *Economic Journal* 109 (459): 598–633.
- Guillaumont, Patrick, and Lisa Chauvet. 1999. "Aid and Performance: An Assessment." Université d'Auvergne, CERDI, Clermont-Ferrand.
- Gunning, Jan Willem. 2000. "The Reform of Aid: Conditionality, Selectivity and Ownership." Paper presented at the Swedish International Development Authority conference on aid and development, 20–21 January, Stockholm.
- Hansen, Henrik, and Finn Tarp. 1999. "The Effectiveness of Foreign Aid and Cross-Country Growth Regressions." University of Copenhagen.
- Haque, Nadeem U., Mark Nelson, and Donald J. Mathieson. 2000. "Rating Africa: The Economic and Political Content of Risk Ratings." In Paul Collier and Catherine Pattillo, eds., *Investment and Risk in Africa*. London: Macmillan.
- IMF (International Monetary Fund). 1997. *The ESAF at Ten Years: Economic Adjustment and Reform in Low-Income Countries*. Washington, D.C.
- Isham, Jonathan, and Daniel Kaufmann. 1999. "The Forgotten Rationale for Policy Reform: The Productivity of Investment Projects." *Quarterly Journal of Economics* 114 (February): 149–84.
- Jaspersen, Frederick Z., Anthony H. Aylward, and A. David Kox. 2000. "Risk and Private Investment: Africa Compared with Other Developing Areas." In Paul Collier and Catherine Pattillo, eds., *Investment and Risk in Africa*. London: Macmillan.
- Kasekende, Louis A., and M. Atungi-Ego. 1999. "Uganda's Experience with Aid." *Journal of African Economies* 8 (4): 617–49.
- Killick, Tony. 1998. *Aid and the Political Economy of Policy Change*. New York: Routledge.
- Lensink, Robert, and Oliver Morrissey. 1999. "Uncertainty of Aid Flows and the Aid-Growth Relationship." University of Nottingham, Centre for Research in Economic Development and International Trade.
- Lensink, Robert, and Howard White. 1999. "Assessing Aid: A Manifesto for Aid in the 21st Century." SIDA Evaluation 1999/17:13. Swedish International Development Authority, Stockholm.

- Oyejide, Ademola, Benno Ndulu, and Jan Willem Gunning. 1999. "Introduction and Overview." In Ademola Oyejide, Benno Ndulu, and Jan Willem Gunning, eds., *Regional Integration and Trade Liberalization in Sub-Saharan Africa. Volume 2: Country Case-Studies*. London: Macmillan.
- Rodrik, Dani. 1989. "Credibility of Trade Reform—Policy Maker's Guide." *World Economy* 12 (March): 1–16.
- . 1996. "Why Is There Multilateral Lending?" In Michael Bruno and Bruno Pleskovic, eds., *Annual World Bank Conference on Development Economics 1995*. Washington, D.C.: World Bank.
- Sachs, Jeffrey D., and Andrew M. Warner. 1995. "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity* 1. Washington, D.C.: Brookings Institution.
- Stiglitz, Joseph E. 1998. "Distribution, Efficiency and Voice: Designing the Second Generation of Reforms." Paper presented at the Brazilian Ministry of Land Reform and World Bank conference on asset distribution, poverty, and economic growth, 14 July, Brasilia, Brazil.
- van Donge, Jan Kees, and Howard White . 1999. "Swedish Programme Aid to Vietnam: An Evaluation." SIDA Evaluation 199/17: 7. Swedish International Development Authority, Stockholm.
- White, Howard. 1999. *Dollars, Dialogue and Development: An Evaluation of Swedish Programme Aid*. Stockholm: Swedish International Development Authority.
- World Bank. 1998. *Assessing Aid: What Works, What Doesn't and Why*. New York: Oxford University Press.

Beyond Rosenstein-Rodan: The Modern Theory of Coordination Problems in Development

Karla Hoff

Coordination problems—and solutions—help determine whether development occurs. More than 50 years ago Paul Rosenstein-Rodan pointed out that spillovers may cause the return to an activity to increase with the number of other actors undertaking the same or complementary activities. If such spillovers are strong enough, multiple equilibria may exist—some of them better for everyone than the alternatives, but with no tendency for market forces to lead to the better state of affairs. Modern economic theory has broadened our view of the sources of spillovers that could lead to underdevelopment as an equilibrium. This article argues for an “ecological” perspective on development, where the influences from others in one’s environment are a critical determinant of outcomes, and many interaction effects are not mediated by markets. This perspective distinguishes between deep interventions, which change underlying forces, and shallow interventions, which do not and may actually make things worse.

Paul Rosenstein-Rodan (1943) famously argued that at an early stage of development, the investments of industrializing firms in one sector may increase the profitability of other sectors throughout the economy. Simultaneous industrialization of many sectors could be profitable for all of them, but no sector would be profitable industrializing alone. As a result, an underdevelopment equilibrium was possible: even the market may not succeed in coordinating the activities needed to ensure development.

In modern terms there could be a *coordination failure*, where individuals’ failure to coordinate complementary changes in their actions leads to a state of affairs that is worse for everyone than some alternative state of affairs that is also an equilibrium.

Karla Hoff is a research economist in the Development Research Group at the World Bank. This article draws on Hoff and Stiglitz (2001a). The author thanks Irma Adelman for helpful comments and to Abhijit Banerjee, Arnold Harberger, Gustav Ranis, Debraj Ray, and Joseph E. Stiglitz for useful discussions of the issues raised in this article. Financial support from the MacArthur Foundation for the research that underlies this article is gratefully acknowledged.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

The obstacle to achieving the better state of affairs is not a matter of technological opportunities (or even knowledge of those opportunities), or of resources or preferences, but only of coordination.

In the 1950s the possibility of “underdevelopment equilibria” and “vicious circles of poverty” was elaborated on by many other authors (including Nurkse 1953; Leibenstein 1957; Myrdal 1957; and Hirschman 1958). But individuals are rarely convinced by those who do not address their concerns. Without a well-developed theory of the sources of spillover effects (externalities), the idea of an underdevelopment equilibrium had little influence on neoclassical economists of that period. They continued to argue that the market could coordinate the changes needed for development (Krugman 1993, 1995; Stiglitz 1993).

Advances in the past 30 years in the modeling of externalities, technological progress, and scale economies have made it possible to provide formal models that capture Rosenstein-Rodan’s insights. At the same time, information economics has shown that neoclassical theory is based on a special case—perfect, costless information—and that its implications for efficiency are no longer valid as soon as one departs in almost any respect from that special case.¹ Before these theoretical developments, most economists thought that the implication of externalities was that the economy would be slightly distorted. But we now understand that the interaction of slightly distorted behaviors of many different agents may produce very large distortions.²

Although it formalizes Rosenstein-Rodan’s basic insight, modern work on coordination problems is far from his original story in several respects:

- *Remedies through international trade.* Rosenstein-Rodan was particularly concerned with the demand effects associated with industries experiencing increasing returns. In some cases this argument loses its force when an economy is opened to international trade, as Tinbergen (1967) pointed out.
- *Other channels of spillovers.* Whereas Rosenstein-Rodan focused on demand effects, modern theory has identified many other channels of spillovers that can lead to coordination failures: spillovers in the technology of an individual agent, spillovers mediated by social or political interactions, information externalities, and spillovers when agents come together through a search process.³ The result can be coordination problems for a wide range of behaviors that matter for development. This article considers examples involving corruption, innovation, contract enforcement, and property rights.
- *Role of government.* Rosenstein-Rodan argued that there could be coordination “from above” when government planned the process of industrialization. In contrast, most recent scholars recognize that there is no such thing because government is part of the endogenous set of institutions to be explained (see Dixit 1996 and Basu 1997, 2000). Moreover, governments fail even in democracies, just as markets do.⁴ But a positive development of recent years, discussed below, is to try more limited interventions to harness the spillovers among agents, and to try to sequence policy reforms in a way that makes it more likely for good equilibria to emerge.

This article shows how developments in economic theory in the past 30 years have broadened our understanding of the channels of spillovers and made possible a new understanding—which I refer to as the “ecological perspective”—of pitfalls and opportunities in development.

The Place of Coordination Failures in Modern Economic Theory

Neoclassical Theory and the Coase Theorem

Neoclassical theory, as I use the term here, was the central economic paradigm until the mid-1970s. It studies price-mediated allocations in a setting of “complete” markets. There is a market for all commodities that are, or could be, produced. “Virtual” prices ensure that actions are taken if their social benefits exceed their opportunity costs.

Such models served as a benchmark. Neoclassical economists recognized that their models did not take into account direct interdependencies among individuals that were not mediated through markets (such as Pigou’s polluting factories). But by and large, neoclassical economists believed that these were relatively unimportant and that interdependencies that were mediated through markets (*pecuniary externalities*) did not have efficiency effects (Mishan 1971; Cheung 1973).⁵

In an extension of neoclassical economics, Coase (1960) argued that when an economy departed from the complete markets assumption of neoclassical theory, private agents would negotiate efficient outcomes (if there were no transaction costs). In that tradition, North and Thomas (1973) argued that the institutional changes in the West between 900 and 1700 were driven by changes in the scarcity values of products and factors. As with competition among firms in neoclassical economics, superior institutions ultimately eliminated inferior institutions.

Institutional Economics outside the “Straitjacket” of Neoclassical Economics

But another strand of the literature recognized that markets were inherently limited by problems of information and enforcement.⁶ There could never be the complete set of markets in goods, risks, and futures on which the efficiency results of neoclassical theory rest. And transaction costs are important, particularly in economic development. Nonmarket institutions arise in response to limits on markets, but there need not be forces that bring about efficiency.

This can be put another way. Individual actions that are privately rational need not be socially rational when the individual takes his environment as fixed. There might be another set of individual actions that would create a different environment in which a different set of actions would be an equilibrium, and that environment might be a better state of affairs. And there could be multiple equilibria in institutions, with no tendency for market forces to select the one that was best.

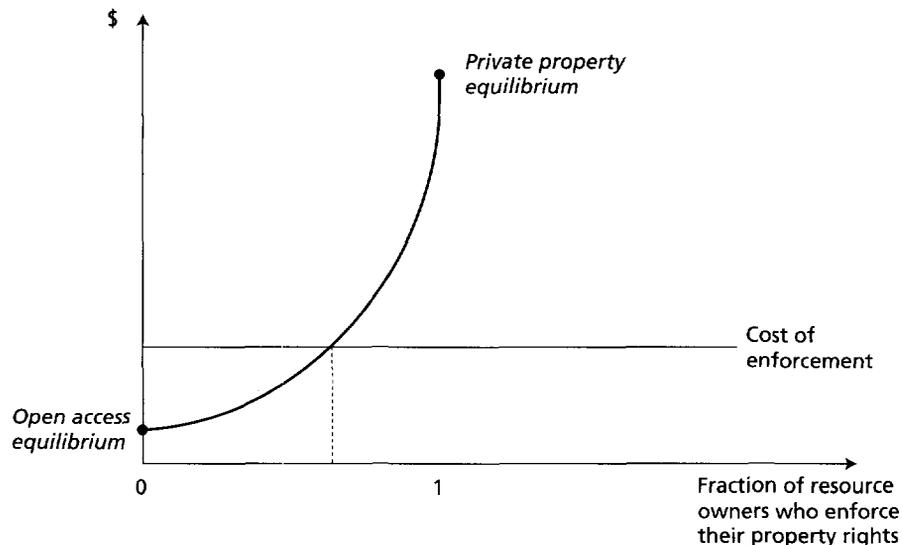
This can be illustrated in a simple way, following de Meza and Gould (1992). Suppose that property rights are well defined but that owners of forested land must incur a fixed cost—say, building a fence or hiring a patrol—to enforce their property rights over the forest. Without this private expenditure, workers will be free to use the private forests and extract income from them. By enforcing their property rights, owners can hire workers to exploit their forests and collect the resource rents.

In deciding whether to enforce their property rights, owners compare the potential rents and the enforcement cost. Rents will be larger if workers' reservation wage is lower. The reservation wage, in turn, depends on how many other owners are enforcing their property rights. As the number of owners who enforce their rights increases, the outside opportunities of workers fall, and so does the reservation wage. With lower wages, potential resource rents rise. Thus two stable equilibria may exist, indicated in figure 1. In the first, all owners enforce their property rights, wages are low, and rents are high. In the second, no one enforces, wages are high, foregone rents are low, and aggregate surplus is maximized.

As Coase (1937) emphasized, when the enforcement of property rights is costly, a market may or may not be the best allocation system. But as those working in the Coase tradition did not generally recognize, whatever the best allocation system is, a decentralized economy may not achieve it.

A general insight of recent theoretical work is that while institutions may be established to improve economic outcomes, there is no assurance that will happen. Institutions may be part of an equilibrium yet be dysfunctional. For example, Arnott and Stiglitz (1991) consider the consequences of the social institutions that arise as a result of the incomplete insurance provided by markets

Figure 1. Multiple Equilibria in the Level of Enforcement



Source: Author's illustration.

because of moral hazard. They show that informal social insurance may crowd out market insurance and lower welfare. More generally, developing countries may be caught in a vicious circle where limited market development results in large information imperfections, and these information imperfections give rise to institutions—such as informal, personalized networks of exchange (Kranton 1996)—that impede the development of markets. Banerjee and Newman (1998) show that an inefficient dualistic economic structure, where market forces govern economic exchange in one sector but not in another, can be explained as the consequence of a self-sustaining network in the informal sector. Dualism may be one equilibrium among several.

A further implication of the modern theory of coordination failures is that improvements in existing institutions—“good mutations”—may not survive if they require accompanying changes in other social institutions. For example, “if the institutional matrix rewards piracy, then [only] piratical organizations will come into existence” (North 1994, p. 361).

History Dependence and Poverty Traps

Recent historical accounts go beyond the observation that there are multiple equilibria and show that economic outcomes exhibit history dependence. History (or path) dependence means not just the obvious fact that past events may circumscribe current choices, but that past events determine and predict the decisions at all crossroads since some initial starting point. For example, after a wartime disruption of trade, an economy may undergo a structural change, develop a new system, and never return to the original one.

The distribution of wealth is one of the most important channels through which history can have potentially large, permanent effects. Here I summarize a few central ideas from a rapidly growing literature. One line of thought emphasizes political influences of wealth. Engerman and Sokoloff (1997) and Sokoloff and Engerman (2000), for example, find that the highly unequal distribution of wealth in the colonies of North and South America that practiced plantation agriculture had long-lasting effects on legal, educational, and political institutions. Another line of thought, discussed in Hoff (1994, 1996), Hoff and Lyon (1995), and Bardhan, Bowles, and Gintis (2000), emphasizes that with imperfect capital markets, poor people have limited access to credit and may be unable to enter into labor and land rental contracts that provide strong incentives for effort. For example, they may be sharecroppers instead of renters, and entrepreneurial activities may be foreclosed. Some initial distributions of wealth are associated with subsequent economic growth, but for others growth paths do not exist.⁷

History also affects outcomes by affecting beliefs. An obvious case is one where expectations are at least partly adaptive: people expect others to behave in the future as they have in the past. But even with fully rational expectations, history can cast a long shadow.

For example, an outbreak of corruption, or the revelation that some firms in an industry passed off shoddy goods as high-quality goods, can tarnish the reputation

of the whole industry and reduce the incentive of every member of the group to behave honestly in the future. Suppose that the reputation of a group member (such as an employee in an organization or a firm in an industry) depends on the member's past behavior and, because the member's track record is observed with noise, on the group's past behavior. As Tirole (1996) shows, the revelation that some group members were dishonest in the past will increase the time it takes for any agent to establish a reputation for honesty. This will reduce his incentives for honesty and may create a vicious circle of corruption: "the new members of an organization may suffer from the original sin of their elders long after the latter are gone" (p. 1).

A Radically Broadened View of Externalities and Public Goods

Coase recognized the limits imposed by transaction costs, but many of his followers did not. They tended to argue (counter to the view in Rosenstein-Rodan) that real-world externalities were hard to find. In the 1970s, the iconic example of externalities was still Meade's apple farmer and beekeeper, who each provide "unpaid factors" to the other. But Cheung (1973) found in the cases he investigated that contracts internalized these externalities. The iconic example of a public good was the lighthouse. But Coase (1974) argued that throughout much of British history, private arrangements had addressed this problem, too.⁸

Modern theory, in contrast, suggests that those working in the Coase tradition were looking for inefficiencies *in the wrong place*. The iconic examples of technological externalities—the pastoral examples and the lighthouse—are ones that contracts among the affected parties, or mergers among firms, may be able to internalize. But many other externalities are too diffuse to be amenable to private solution. They include information externalities, group reputation effects, agglomeration effects, and knowledge spillovers. They also include pecuniary externalities (Newbery and Stiglitz 1982, Greenwald and Stiglitz 1986). Pecuniary externalities often look analytically like externalities of the familiar technological sort. Recall the example above on property rights, where as more owners fenced in their forests, the return to fencing went up. In that example the spillover effects were transmitted only through the change in the wage rate.

Modern theory radically broadens our notion of spillover effects. It shows that in many cases even the set of equilibrium prices has public good properties. When spillover effects are sufficiently strong, there can exist multiple, Pareto-ranked equilibrium outcomes—each supported by some set of prices. Each individual's choices contribute to the selection of one of those equilibria, but he ignores his effect on the outcome.

"Ecological Economics"

Whereas neoclassical economics emphasizes the forces pulling toward equilibrium—and with similar forces working in all economies, all should be pulled toward the

same equilibrium—modern development economics focuses more on evolutionary processes, complex systems, and chance events that may cause systems to diverge. Thus modern theory tends to be influenced more by biological than physical models. Near the end of *The Origin of Species*, Darwin (1993 [1859]) wrote, in thinking about the Galapagos Islands:

[The plants and animals of the Galapagos differ radically among islands that have] the same geological nature, the same height, climate, etc. . . . This long appeared to me a great difficulty, but it arises in chief part from the deeply seated error of considering the physical conditions of a country as the most important for its inhabitants; whereas it cannot, I think, be disputed that the nature of the other inhabitants, with which each has to compete, is at least as important, and generally a far more important element of success. (p. 540)⁹

The economy is like an ecosystem, and Darwin was implicitly recognizing that ecosystems have multiple equilibria. Far more important in determining the evolution of the system than the fundamentals (weather, geography) are the endogenous variables, the ecological environment. Luck—accidents of history—may play a role in determining that and, thus, in the selection of the equilibrium.

The main differences between the old and new views of spillover effects are summarized in table 1. From these effects it is easy to formulate models with underdevelopment equilibria, as the next section illustrates.

Coordination Problems in Development

This section presents an overview of theories of development in which spillover effects exist and are so large that some societies may be characterized by a high

Table 1. Old and New Views of Externalities and Public Goods

	<i>Old view: Iconic examples</i>	<i>"Ecological" view</i>
Externalities	Beekeepers and apple farmers Polluting factories	Pecuniary externalities arising from: <ul style="list-style-type: none"> • Enforcement of property rights • Information spillovers • Ownership structures • Demand effects on nontradables produced with increasing returns Knowledge spillovers Externalities in contract enforcement Search externalities Social and political economy interactions The equilibrium set of prices
Public goods	Lighthouses	Group reputation Knowledge (say, that a certain technological result is feasible)

level of innovativeness, whereas others—which are intrinsically similar—are characterized by a low level. In these theories, an individual's behavior depends on the environment, and among the most important elements of the environment are the behaviors of others. If the incentive for agents to choose an action is sufficiently positively related to the number of agents who choose similar actions, then an economy can have multiple equilibria, some better for every agent than an alternative equilibrium.

Knowledge Spillovers

One way to conceptualize the research and development (R&D) that producers undertake is that it transforms known facts and accepted principles into potentially profitable new applications. In this view, the expected return to an investment in R&D increases with the stock of ideas in the public domain. If part of the outcome of private research seeps into the local public pool of knowledge, then the more research that is conducted, the larger the pool of ideas on which each producer in the community draws. (There is persuasive evidence that knowledge spillovers are strongest over short distances; see Jaffe, Trajtenberg, and Henderson 1993.) Thus one possible impact of expanded R&D is a greater incentive to invest in R&D (Romer 1986).

To analyze this situation, consider a simple model with a large number of identical producers. The profit (utility) of any producer depends on prices, the producer's level of R&D (denoted by a^i , which can be any value between 0 and 1), and the level of R&D of all others. Because I will focus on symmetric equilibria, I consider the case where all other producers choose the same level of R&D (denoted by a). Thus the profit function is written as

$$u^i = u^i[a^i; a, p(a)]$$

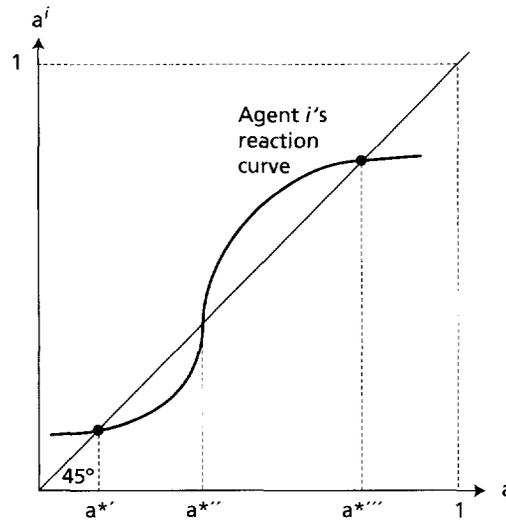
where p is the price vector (which depends on the agents' actions). Assume for each agent diminishing marginal returns to his action. Each agent chooses a level of action to maximize his utility given the actions of others. (Each agent is small enough that he ignores the direct effect of his actions on others and on p .)

Let u^i_1 represent the partial derivative of u^i with respect to the first argument (a^i). The reaction function,

$$u^i_1 [a^i; a, p(a)] = 0$$

characterizes the action that the representative agent i will take for all possible values of a selected by the remaining actors. The reaction function states that, given a , the agent cannot obtain a higher payoff by a marginal change in the level of his action. Figure 2 depicts the case where a higher action a by all other agents *will lead the remaining agent i to follow suit*. A higher action by other agents increases the marginal return to higher action by each. The actions of different agents are *strategic complements*.

Figure 2. Multiple Equilibria in a Model with Symmetric Agents



Source: Author's illustration.

The interior, symmetric equilibria are values of a^* that solve the equation

$$w_1^i [a^* ; a^*, p(a^*)] = 0.$$

Figure 2 illustrates a stable low-level equilibrium at $a^{*'}$ and a stable high-level equilibrium at $a^{*'''}$.¹⁰ By themselves, equilibrium prices do not provide the incentive needed to implement the efficient equilibrium ($a^{*''}$).

Externalities from Innovative Behavior

In some countries, individuals appear to be constantly looking around for new ideas. In other countries, they appear to resist them. These behaviors create externalities; each individual's choices may thus depend on the choices made by other people. Sah and Stiglitz (1989) formalize this idea. They construct a model of social equilibrium in which individuals can choose to behave bureaucratically or innovatively. Bureaucrats make life more difficult for innovators, and vice versa. Let x be the share of the population that chooses to be innovative. Let $U(I; x, p)$ be the utility associated with the innovative strategy, and let $U(B; x, p)$ be the utility associated with the bureaucratic strategy. Each individual chooses the activity that yields the greater utility, taking x and p as given, where p is a function of x : $p = p(x)$. Then it is easy to see that there may be multiple equilibria. If most people choose to behave bureaucratically, it may pay only a few people to behave innovatively, and vice versa. An interior equilibrium (where x is between 0 and 1) is a fraction x^* that solves the equation

$$U[I; x^*, p(x^*)] = U[B; x^*, p(x^*)].$$

Equilibria also may exist where all agents make the same choice, one entailing bureaucratic behavior if

$$U[I; 0, p(0)] < U[B; 0, p(0)]$$

and another entailing innovative behavior if

$$U[I; 1, p(1)] > U[B; 1, p(1)].$$

A slight variant of this model can be used to explore evolutionary dynamics. Assume that rather than the individual choosing to be innovative or bureaucratic, each individual has a certain fixed type that may reflect his upbringing, and agents live just one period. Different reproductive rates are a function of utility levels so that

$$d \ln x/dt = k [U(I; x, p) - U(B; x, p)]$$

for some positive constant k and time variable t . Then the set of equilibria will be the same as before, and the equilibrium on which the economy converges will depend on its history. Historical events—such as a country's opening to international competition that hurts bureaucratic firms—may move the economy from one equilibrium to the other, thereby affecting the long-run rate of technological progress.¹¹

Externalities from Rent-Seekers

A variant of the preceding model focuses on externalities from rent-seeking. Rent-seeking is broadly defined as unproductive activities that use resources and extract income from others—such as taking bribes, lobbying for preferences, and predation. For example, Murphy, Shleifer, and Vishny (1993) consider a farm economy where each individual chooses to engage in one of three activities:

- Producing a cash crop for the market, but with output from the activity vulnerable to rent-seeking.
- Producing a subsistence crop, with output not vulnerable to rent-seeking.
- Being a rent-seeker, and expropriating part of the output of cash crop producers.

An equilibrium is an allocation of the population among the three activities. Suppose that over some range, as more resources move into rent-seeking, returns to cash crop production fall faster than returns to rent-seeking. As a result, the returns to rent-seeking increase relative to production for the market. This can give rise to multiple equilibria. An economy can be in an equilibrium where the number of cash crop producers is low and returns to such production are low because the number of rent-seekers is high. But there is another equilibrium where the opposite is true. As Acemoglu (1995) emphasizes, this means that economies with the same initial opportunities may diverge over time—some towards a high growth path where

activities that increase aggregate output have high economic and social rewards, and others towards a low-growth path where they do not.

Externalities from Enforcement Mechanisms

Most development economists agree that among the most important institutions in an economy are those that enforce contracts. Greif (1994) examines cultural factors that might explain why two premodern societies—the Magribi in North Africa and the Genoese in Italy—evolved along different trajectories of social organization. To illustrate the main ideas, he presents a model with two kinds of actors: merchants and agents. Agents carry out overseas trade on behalf of the merchants and choose between two strategies: to be honest or to cheat. A merchant also chooses between two strategies: to rely on collectivist or individualist means to punish cheaters. Under *collectivist enforcement*, a merchant refuses to hire an agent known to have cheated any merchant in the collective group. Under *individualist enforcement*, a merchant punishes only agents who have cheated him. Greif shows that if a merchant believes that collectivist enforcement is likely, then it may not be in his interest to hire an agent known to have cheated other merchants. That makes such expectations self-fulfilling.

The intuition for this result is straightforward: an agent who has already damaged his reputation has little to lose by cheating again, and so will be more easily tempted to cheat his current employer than would an agent with an unblemished reputation. That makes the agent who has damaged his reputation an undesirable hiring prospect. But if the merchant believes that individualist enforcement will occur, the motive for collectivist enforcement is absent. Thus two equilibria may exist, one entailing collectivist enforcement and the other individualist enforcement. The equilibrium that is “selected” will depend on beliefs (culture).

In the short run, individualist enforcement will be more costly because it forgoes the stronger, group-level punishment mechanism. But in the long run, individualist enforcement will increase the likelihood that government takes on a role in enforcing contracts. By widening the scope of markets through which anonymous trading can occur, such institutions promote long-run growth. Greif (1994) interprets the history of the West in just such terms.

Local Externalities from Community Stakeholders

Civil society is an important vehicle for change in institutions and policies. Community participation is also a factor in the success of development projects (Isham, Narayan, and Pritchett 1995). There is persuasive evidence from developed countries that the individuals who are most likely to participate in community associations, respond to neighborhood problems, vote, and maintain their property are those with a secure property stake in their residence.¹² Since stakeholders create local spillovers, households’ decisions over stakeholding are interdependent. Hoff and Sen (2000) present a framework in which to examine the possible consequences of local spillovers from homeownership.¹³ They show that even if all households have identical preferences and

identical abilities to contribute to their communities, polarized cities can emerge—the rich in homeownership communities with high levels of civic engagement and home improvement effort, the poor in rental communities with low levels of both.

The basic story goes as follows. Every household desires to own its home because it realizes that then it will obtain a greater return on any effort it makes to improve its home and community (since it can “capitalize” the gains from these investments). However, because of capital market imperfections, a household can afford to buy its home only if its income is above a threshold. So the richer households buy their homes and expend more effort in improving their homes and communities than do the poorer households, who rent. It is in these effort choices that there exist community-wide interaction effects. The quality of a home and its future market value are higher when it is in a community where neighbors expend high effort (a spillover effect); and for an individual household, the marginal returns to effort are higher when neighbors expend high effort (a complementarity effect).

The interplay of these effects influences community formation. Compared to renters, homeowners take more out of the community and so may be willing to pay more for land in a community with a larger number of stakeholders. This case is depicted in figure 3A. A typical indifference curve of a homeowner (denoted by \bar{U}^e because he expends effort in home and community improvement) is steeper than the indifference curve of a renter (denoted by \bar{U}^n , because a renter expends negligible effort). Homeowners segregate in response to this complementarity relation, no matter how weak. (There cannot be a stable equilibrium where renters and homeowners live together in more than one community, for the homeowners would always be willing to bid more for a home in the high-home-owning community than the renter would. Homeowners and renters would continue to switch places until at least one community was completely segregated by tenure.) The outcome can be socially very inefficient. Figure 3B depicts the case where segregation of a community that was initially 50 percent homeowner results in a large loss in housing quality to renters that more than offsets homeowners’ small gains in housing quality. But land price differentials support the segregated equilibrium.

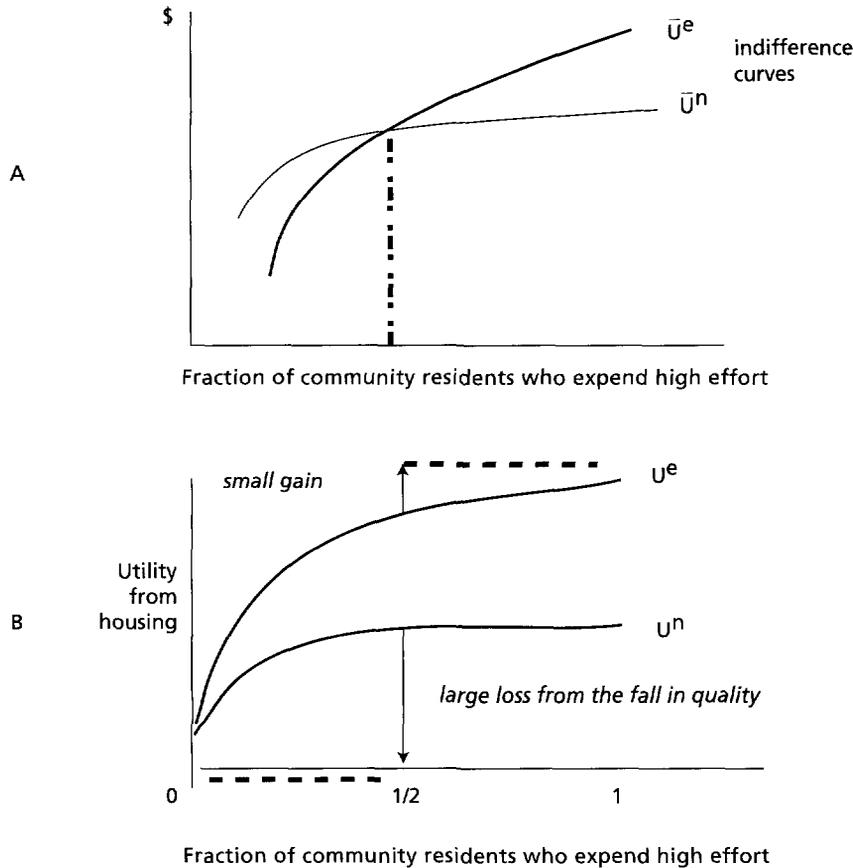
To summarize, interaction effects across community residents may cause households to self-organize into “good” and “bad” communities. The capital market frictions faced particularly by the poorer households induce in them choices over contracts and behaviors that make them less desirable as neighbors. This, in turn, induces a pattern of community formation where “bad behavior” is concentrated. The implication is that market forces together with social interaction effects can produce a class of communities that, for lack of local stakeholders, cannot respond to development opportunities.

“Big Push” Theories of Industrialization

Let us now return to the central concern of Rosenstein-Rodan: that an economy with all the preconditions for industrialization would fail to industrialize because of a failure to coordinate complementary investments. The relevance of models of

Figure 3. Stakeholding, Local Interactions, and Segregation

(A) Compared to renters, homeowners take more out of the community and so, as depicted in Panel A, they may be willing to pay more for land in a community with a larger fraction of residents who expend high effort. (B) This will lead to tenure-segregated communities, which may be socially very inefficient, as Panel B illustrates.



Source: Author's illustration.

coordination failures rests on diffuse externalities. If externalities were not diffuse, negotiation among the affected parties—or mergers among firms—should internalize them. Rosenstein-Rodan was particularly concerned with diffuse externalities through demand effects associated with industries experiencing increasing returns. But if all goods are tradable, this argument loses its force.

Thus the modern theory of underdevelopment equilibria based on demand effects focuses on nontradable inputs into production. Complicated technologies often require a variety of local inputs and producer services. Increasing returns in their production can generate external economies at the level of final (tradable) goods. An expansion of industry in an economy increases the demand for these nontradable

inputs, which lowers their costs and increases the available variety. With greater variety of intermediate inputs, production of final goods may be more efficient (for example, computer technicians who specialize in certain programs can troubleshoot those problems faster than can generalists). Thus it can be the case that when all other sectors industrialize, it pays the remaining sector to do so. But when all other sectors use a traditional technology that does not require intermediate inputs, it pays the remaining sector to do so too. An underdevelopment equilibrium can thus be sustained even when the economy is fully open to international trade (Helpman and Krugman 1985, ch. 11; Rodríguez-Clare 1996; Rodrik 1996).

There are a variety of ways to think of the nontradable inputs. One is that they represent different categories of specialized skilled labor, such as computer technicians and software designers. As Rodrik (1996, p. 2) argues, "A worker's decision to invest in a specialized skill depends both on the demand for the particular skill and the existence of complementary skills in the economy." But why can't a single firm train the labor force it needs and so internalize the externalities among the decisions of employees and between their decisions and those of the firm? Acemoglu (1997) suggests one reason. He shows that even perfect contracting within a firm and a perfect capital market may fail to internalize the social consequences of the decisions made by workers and firms. His explanation relies on search costs.

Suppose that there are two kinds of actors: firms that choose whether to adopt a new technology, and workers who choose whether to become trained to use the new technology. Training pays off only if the worker is employed by a firm that has innovated, and innovation pays off only if the firm employs a worker who has trained. There is a large and equal number of firms and workers. Each firm employs one worker and there are two time periods. In the first period, a firm is matched with one worker, they jointly make decisions on training and innovation, and there is complete contracting between them (that is, no information problems or transaction costs). At the end of the first period there is some risk of separation. If separation occurs, a firm has to find a new worker, and a worker has to find a new firm. In the second period, output is produced.

If there were no search costs in the labor market, separation between a firm and a worker would not create a loss. If separation occurred, the worker would simply move to another firm that had adopted the new technology, and all the surplus from training and investment would be captured by the firms and workers who made the investments. But suppose that search is costly. With costly search, matching will be imperfect. There is no guarantee that the firm with the investment in the new technology will be matched with the worker who has the training.

Multiple equilibria can occur because a firm's likelihood of finding the right worker depends on the thickness of the market (the number of trained workers). The worker's likelihood of finding the right employer also depends on the thickness of the market (the number of firms that have adopted the new technology). Without a risk of separation, there would be no inefficiencies because there would be no interactions with future employees or employers. The inefficiency arises because of an externality between workers and their future employers, and between firms and

their future employees, that cannot be internalized because the identity of the actor with whom one may be matched is unknown.

In contrast, Rosenstein-Rodan identified the obstacles to training with incomplete contracting between a given firm and its employee: “There are no mortgages on workers—an entrepreneur who invests in training workers may lose capital if these workers contract with another firm” (1943, p. 205). This imperfection is not necessary to generate inefficiency. Search costs create the possibility of multiple, Pareto-ranked equilibria, where some countries are characterized by a high level of innovativeness, and other countries by a low level. Training and innovation within any single firm may pay off only if a sufficiently large fraction of the population trains and innovates.

An Econometric Test of Spillovers and Local Poverty Traps in Rural China

Testing has become possible with the formalization of coordination problems and underdevelopment traps. It is beyond the scope of this article to review the empirical literature, which is still at an early stage.¹⁴ This section presents a promising line of investigation based on work using Chinese census data by Jalan and Ravallion (1998) and Ravallion and Jalan (1996, 1999).

China is a good place to examine the theory for several reasons. Because the government severely limits geographic mobility in China, and because capital mobility is also limited, factors move little in response to different opportunities across regions. Thus it may not be too unrealistic to treat the assignment of households to particular counties as random (and not the result of self-selection) and to view each household’s investment opportunities as a function of its local opportunities. Further, China exhibits remarkable regional differences in living standards and growth rates. For example, in 1990 rural poverty in the inland mountainous province of Guizhou was 7–10 times (depending on the poverty line) that in the neighboring coastal province of Guangdong—just a few hundred kilometers away.

The census data come from a 1985–90 panel of 5,600 farm households in southern China. Jalan and Ravallion estimate two models derived from optimizing behavior.¹⁵ The first is a simple expository model with just two explanatory variables for growth in household consumption (apart from household time-varying fixed effects): initial household wealth per capita (*HW*) and mean wealth per capita in the county of residence (*CW*). The equation for the growth rate, *g*, is (t-ratios in parentheses):

$$g = -0.143 - 0.0166 \ln HW + 0.0378 \ln CW$$

(5.61) (5.91) (8.13)

Counties with higher average wealth showed higher average growth in household consumption. Note also that the sum of the coefficients on *ln HW* and *ln CW* is positive, which implies that the effects are large enough to lead to divergence across counties, which is the aggregate pattern over the six-year period of the survey. Jalan and Ravallion interpret their results to mean that an increase in a county’s average wealth increases the marginal return to household wealth. This is due entirely to

geographic externalities rather than to increasing returns to wealth at the household level, since the negative coefficient on initial household wealth ($\ln HW$) implies that there are decreasing returns to capital within households.

I illustrate these results in figure 4. Let K denote the household's capital stock and \bar{K} denote the average household capital stock in the county. The econometric results suggest that the marginal productivity function is downward sloping with respect to household capital, and shifts up with an increase in the county's average level of capital. (Ignore points 1 and 2 in the figure for now.)

Figure 5 depicts the equation $g = 0$; that is:

$$g = -0.143 - 0.0166 \ln HW + 0.0378 \ln CW = 0$$

and plots the values of household wealth and county wealth from the survey. Note that every household is near the critical line. (Readers can check their understanding of the figure by noting that a typical household whose per capita wealth is $\ln 6$ yuan will enjoy rising or falling consumption over time as the household lives in a county with per capita wealth higher or lower than $\ln 6.5$ yuan.) For a large subset of the data county wealth is too low, given household wealth, to permit rising consumption. Spillover effects appear to be large enough to generate poverty traps.

Looking back at figure 4 makes it clear why a household in a poor county might have less incentive to accumulate capital than a household in a richer county. Suppose that all households in a county are identical, so $K = \bar{K}$. At point 1, household capital is low and the returns to new investment are low because other households' capital is low. But there is another equilibrium at point 2 where the reverse is true.

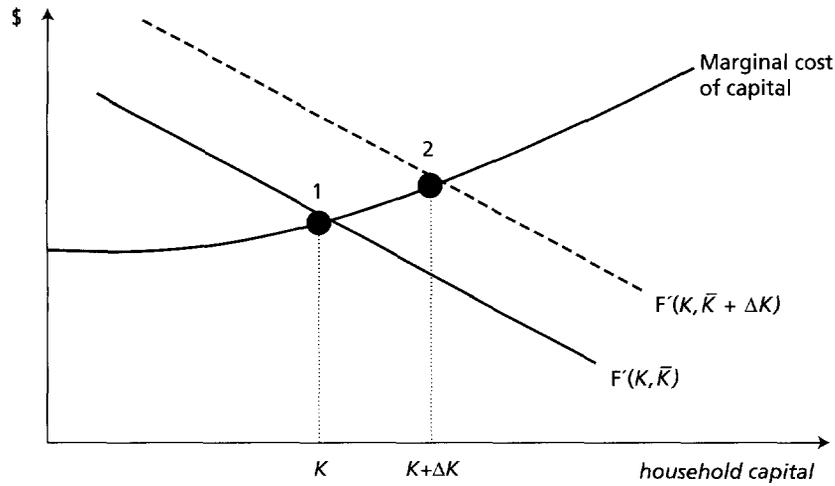
The second model that Jalan and Ravallion estimate does not use county wealth but instead the detailed county variables listed in table 2. The table shows the main results. (Dummies for county, period, mountains or plains, coast or inland, and the like have been deleted. All the regressions were run with household time-varying fixed effects.)

County-level variables related to agricultural modernization (farm machinery use per capita and fertilizer use per cultivated area) and to the share of the population employed in nonfarm commerce had highly significant positive effects on individual consumption growth rates. Holding all else constant, a one standard deviation increase in farm machinery use in an area adds 0.6 percentage points to annual consumption growth, and a one standard deviation increase in fertilizer use adds 1.5 points. By comparison, a one standard deviation increase in rural road density adds 0.7 points to annual consumption growth. Following the same basic procedure as above (setting $g = 0$ and evaluating all but one of the variables at its mean value), Jalan and Ravallion show that these spillovers are big enough to generate underdevelopment traps, consistent with the implication of their simple expository model.

Perspectives on Policy

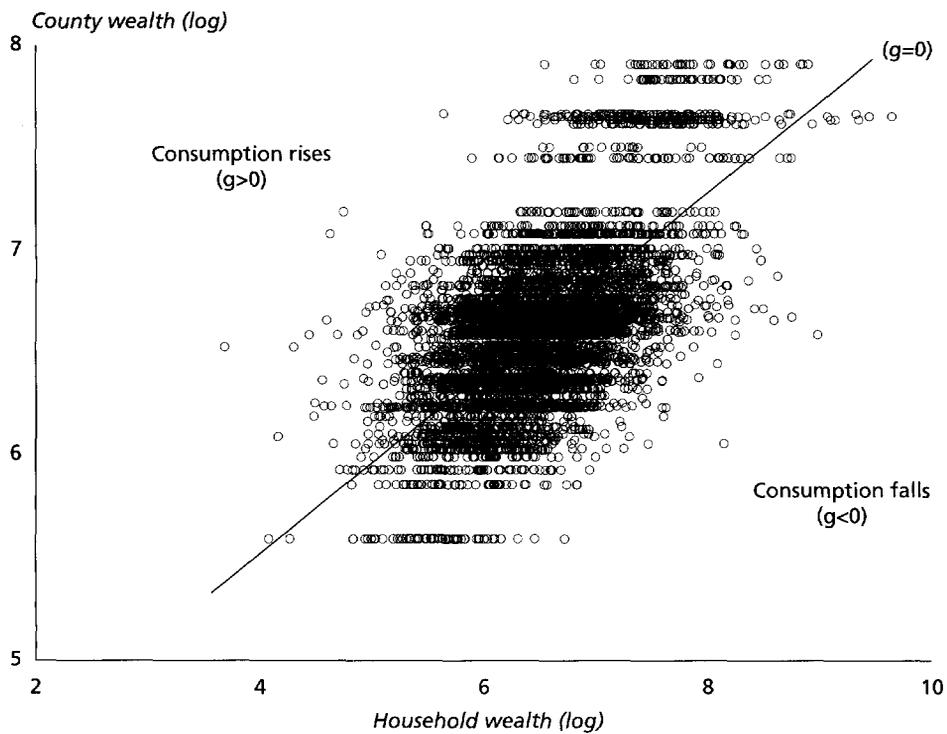
The literature on coordination problems and poverty traps suggests that development may be both easier and harder than was previously thought. Under the older

Figure 4. Marginal Productivity of Household Capital as a Decreasing Function of Household Capital and an Increasing Function of Average County Capital



Source: Author's illustration.

Figure 5. Minimum Levels of County Wealth Needed to Ensure Rising Household Consumption Given Household Wealth



Note: Wealth is measured in yuan per capita at 1985 prices.
 Source: Jalan and Ravallion 1998, figure1.

Table 2. Determinants of Local Opportunities in Rural China

Geographic variable	Coefficient estimate	t-ratio
Farm machinery use per capita (x 1,000)	0.0420 *	3.4328
Cultivated area per 10,000 people	0.0013	1.5013
Fertilizer used per cultivated area	0.0023 *	4.5678
Population density (log)	0.0160	1.6949
Illiteracy rate among people 15 and older (x 100)	0.0159	0.9000
Infant mortality rate (x 100)	-0.0313 *	-2.5295
Medical personnel per capita	0.0011 *	3.6882
Share of population employed in nonfarm commerce (x 100)	0.0067 *	2.1156
Kilometers of roads per capita (x 100)	0.0741 *	4.3033
Share of county population living in urban areas	-0.0228	-1.0254

* Significant at .05 level.

Source: Jalan and Ravallion 1998, table 2.

theory “all” one had to do to ensure development was transfer enough capital and remove government-imposed distortions. Under the new theories “all” one has to do is induce a movement out of the old equilibrium, sufficiently far and in the right direction that the economy will be drawn to a good equilibrium. While this may require fewer resources, it may require more skill. Some policies could lead the economy to a worse equilibrium. That is, even after the policies were removed, their ill effects would persist. And just as the equilibrium set of behaviors in a decentralized economy may not be Pareto efficient, one cannot assume that Pareto improvements are likely to emerge from the political process (see below).

One response to the modern literature on coordination problems has been to emphasize our limited knowledge of the activities that should be coordinated—and hence, the dangers of relying on the state to establish a coordination mechanism (Matsuyama 1995). The problem with highly centralized coordination of activities by government, as occurred in centrally planned economies, is not so much that they never experienced rapid growth but that they suffered from a lack of inventiveness and became “prematurely gray” (attributed to Jeffrey Sachs by Matsuyama 1995). For Matsuyama the problem of coordination is like “*the problem of hundreds of people, scattered in a dense, foggy forest, trying to locate one another*” (p. 134, emphasis added). Neither governments alone nor markets alone can solve it. Early writers who pointed out the right sources of coordination failures drew the wrong policy conclusions when they interpreted coordination failures as a call for a “big push” industrialization centrally planned by the state.

Another strand of the literature emphasizes the incompleteness of models of coordination failures. The selection of the equilibrium is determined outside the model by initial conditions. Making the analysis of intervention precise requires a dynamic framework. Only in a dynamic framework can one ask whether an initial coordination failure will transmit itself over time. Why wouldn’t forward-looking agents with sufficiently low discount rates adopt a path (which might include the option of changing their behavior several times) that would permit as an equilibrium a self-fulfilling move from a bad to a good equilibrium? Is there really any scope for policy?

Adsera and Ray (1998) address these questions in a setting where each agent chooses between two activities (which could be interpreted as entry into a high-technology sector that offers a high return in the long run if it obtains a critical mass, relative to the returns in a low-tech sector). They obtain a striking result: if the positive externalities from moving to the more favorable activities appear with a lag (which can be made arbitrarily short), then the final outcome depends entirely on initial conditions unless there is some gain to being the first to switch. Put another way, without some gain to being among the first to switch, each individual will rationally wait for others to switch first—so no one will switch at all! Thus initial conditions will determine the entire equilibrium outcome.

Interventions to Solve a Coordination Problem

Adsera and Ray show that there is a potential role for policy to enable an economy to break free of history. A temporary intervention can force an equilibrium—but once the equilibrium is attained, the intervention is no longer needed to support it. (This has the advantage that by the time agents have learned how to corrupt the rule-administering process, the rule may no longer be needed.) I consider four examples below.

ANTICORRUPTION PROGRAMS. As Tirole's (1996) model of group reputation, discussed above, suggests, it may be possible for a temporary anticorruption program to switch an economy from a high-corruption equilibrium (sustained by expectations of high corruption) to a low-corruption equilibrium (sustained by expectations of low corruption). An interesting historical case is the Republic of Korea in 1961. Park Chung Hee, president of South Korea from 1961 to 1979, inherited from Syngman Rhee a corrupt bureaucracy that chose policies based on self-enrichment. Within a month of seizing office, Park dismissed the top 10 percent of bureaucrats, jailed a number of the country's leading businessmen for corruption, and sent the rest of the bureaucracy to two-week training courses in management, efficiency, and public spiritedness. He then monitored the performance of economic bureaucrats and quickly shifted them from one bureau to another so that they could not develop corruption networks. These efforts transformed the government's functioning and shifted from soft to hard its interaction with the chaebols (conglomerates). Park then designed the development plan (Mason 1980).

AFFIRMATIVE ACTION. A change in a legal statute may be able to force a new equilibrium if the path to it involves revising beliefs and the revised beliefs sustain the new equilibrium. Stiglitz (1974) shows that affirmative action programs can eliminate equilibria in which productivity is unequal between groups (such as races or ethnic groups) with the same innate abilities but different histories. Productivity between groups may be unequal if, for example, individuals' preferences for education depend on their parents' education and if the resulting differences in education lead to different expectations by employers about the payoffs to training workers. Decisions by employers may then lock groups into different positions in the income

distribution. An affirmative action program changes the behavior of employers, the new behavior creates a new “history” and reveals information about the group that suffered from discrimination, and the revised information can lead to an equilibrium in which prospective employers no longer want to discriminate.

LAW BASED ON SOCIAL NORMS. The law can also serve as a coordination device by changing the extent to which individuals impose social sanctions on violators of a norm. Following Cooter (2000), suppose that a person who punishes someone for violating a social norm risks confrontation or revenge, but that this risk falls as the number of people willing to punish increases. Suppose too that enactment into statutory law of the social norm—say, to use generally accepted accounting standards—lowers the private cost of enforcement because it leads an individual to believe that other individuals will be more likely to impose social sanctions on those who violate the norm. That expectation can be self-fulfilling. There can be multiple equilibrium levels of private enforcement of social norms. Making the social norm into law can “pull in” private activity rather than crowding it out (as occurs in traditional analysis of government provision of public goods), and thus shift society from an equilibrium with little private enforcement of the norm to one with high private enforcement.

By extension, Cooter (1997) argues that a state governed by laws that mirror social norms (a rule of law state) tends to be hard to corrupt—whereas a state where laws are imposed and enforced from above (the rule-of-state law) tends to be costly, ineffective, and easily corrupted. A related view of statutory law comes from Basu (2000), who argues that the only way government-enacted law can influence an economy is to switch it from one equilibrium to another: if an outcome is not an equilibrium candidate absent the law, it is still not an equilibrium under any conceivable legal regime.

RESOURCE STOCKPILING TO DETER POACHING. It is sometimes suggested that the literature on coordination problems implies that progress is possible only if one changes a broad array of policies at the same time. But that is not the general message of the literature. The following example shows how a modest policy intervention—one well within the reach of most developing countries—can resolve a coordination problem in the exploitation of a storable open access resource, such as ivory from elephants.

Kremer and Morcom (2000) consider a setting where more intensive harvesting of open access resources leads to anticipated future scarcity of such resources, which leads to higher current prices, and so to more intensive current exploitation. In particular, elephant poaching can lead to expectations of ivory shortages and thus raise future ivory prices. Because ivory can be stored, the rise in future prices raises current prices, which increases incentives for poaching today. There may be multiple rational expectations of paths of ivory prices and of the elephant population. The paths along which extinction occurs are worse for everyone (not to mention the elephants). Government can eliminate the extinction equilibria by accumulating a sufficient stock of ivory—perhaps from sick elephants—and threatening to flood the market if the elephant population becomes dangerously low. That would drive down the price for ivory and

discourage poaching. In this way, stockpiling policy can change the anticipated price path of ivory and eliminate the extinction equilibria.

A Word of Caution: Deep versus Shallow Interventions

Public choice theory has provided considerable insights into political processes, including the problems associated with the formation of interest groups (Olson 1965, Becker 1983). For instance, free-rider problems help determine which interest groups form, just as they influence the provision of public goods more generally. Public actions affect the costs and benefits associated with interest group formation. Because the costs of forming an interest group are at least partly fixed, interventions that affect the dynamics of the political process—thereby affecting subsequent outcomes—can be thought of as “deep” interventions. They entail irreversibilities.

An example of the dynamics of the political process shows what I mean. Assume that the government is contemplating privatizing a monopoly. There are several potential buyers. All of them have an interest in ensuring that the regulations that prevail after privatization will allow them to enjoy the monopoly profits and perhaps even leverage the monopoly power further. But each, thinking that he has a small chance of winning, is unwilling to spend much to ensure this “collective” good (or bad, depending on one’s perspective). Moreover, each may face large costs in identifying the other potential buyers. Even if a potential buyer identifies the others, if they are numerous there would still be a free-rider problem, with each claiming that they will obtain high profits by increasing efficiency rather than by exploiting monopoly power.¹⁶

But once the privatization has occurred, a single party is the winner. There is no longer a collective action problem, and the winner has the incentive and resources to fight legislation imposing regulation or competition. Thus before the privatization it may be possible to pass rules that promote competition (since there is no organized resistance in the private sector) and there may be public interest groups pushing for it. The sequencing of reforms—that is, whether regulation precedes or follows privatization—matters. In one sequence, the result may be a competitive or regulated industry, with privatization leading to lower consumer prices. In another sequence, the result may be an unregulated monopoly. And the unregulated monopoly may be more efficient than it was as a public sector producer not only in producing goods, but also in exploiting consumers.

It is because history matters that interventions can be effective in the long run. A change (perturbation) to the system on one date can have permanent effects. By contrast, in neoclassical and related theories it is fundamentals—including those associated with the political process—that determine long-run outcomes.

Deep interventions need to be distinguished from shallow interventions that do not entail irreversibility. A deep reform effectively ties the hands of all parties by imposing a high fixed cost of undoing it. A shallow reform does not entail commitment and can make matters worse. Consider the unsuccessful agricultural privatization in Russia in 1991–99 (Amelina forthcoming). In 1991 Russia legalized individual farms and dismantled federal subsidies for cooperative farms. Every farmer gained the right to

become a private producer. Many analysts expected these farmers to exit the inefficient collective system in droves and produce for themselves. Yet the share of agricultural land used by cooperatives fell little: from 91 percent in 1991 to 80 percent in 1997.

When the Soviet system collapsed and cooperatives were left without federal support, district politicians had the incentive and the power to ensure a supply of inputs and subsidies to the cooperatives, just as the Soviet center had done.¹⁷ This support inhibited the potential demand for new market institutions and for suppliers capable of serving small producers. By 1997, there were still no market institutions providing farmers with storage, processing, transport, and insurance (and given the fixed costs of setting up those services, they could succeed only if the number of private farmers was high enough). But there had emerged district-level programs to sustain cooperative farms. Amelina (forthcoming) interviewed farmers to learn why they stayed within the cooperative system. Among the top two reasons in both districts she studied was that “there is no other place to work.”

The Rule of Law in the Transition Economies

Consider again privatization in the transition economies. Only under the rule of law can capitalism fully take root. But how does one establish the rule of law? Recent work suggests that coordination failures can block establishment of the rule of law. One coordination problem relates to financing the legal system (Johnson, Kaufmann, and Shleifer 1998; Roland and Verdier 1999). Another, perhaps more fundamental, coordination problem relates to political support for the rule of law. I discuss this problem here.

Many economists believed that transferring state property to private agents would automatically create strong incentives for investment and support for the rule of law. But this claim overlooked an alternative activity that might also be profitable in an environment where property rights were ill-defined—stripping assets.

Consider two contrasting images of the agents who received control of state assets after privatization:

Privatization offers an enormous political benefit for the creation of institutions supporting private property because it creates the very private owners who then begin lobbying the government . . . for institutions that support property rights. (Shleifer and Vishny 1998, p. 10)

[T]he efficient (responsible) owner . . . became one of the mythological figures of many official programs. . . . [The owner of a privatized firm] doesn't pay wages to the employees, doesn't pay taxes, is not interested in the enterprise's development, establishes subsidiaries in order to “pump out” the assets while leaving only the legal shell of the company, etc. (Radygin 1999, pp. 32–33).

Here I will sketch a model, based on Hoff and Stiglitz (2001b), that argues that the entrepreneur who builds up the value of his firm is not a mythological figure,

but rather a creature of a particular environment. He reflects the opportunities of his environment. As emphasized throughout this paper, an important part of that environment is the behavior of others. As in the case of enclosures of forests, there may be many property rights systems that are equilibria.

The easiest way to think about this problem is through a model that has a very limited set of issues associated with the definition of property rights. Consider a model where the agents are individuals with control rights over privatized property. And the conceivable set of actions is building value (entrepreneurship) or stripping assets, tunneling value out of the firm, and letting the capital stock wear out (stripping, for short). Entrepreneurs support reforms that lead to the rule of law, while asset strippers do not.¹⁹ The rule of law in this case is the enforcement of ownership rights (the most elemental property right). The probability π that the rule of law is enacted depends on the number of entrepreneurs. We assume that this probability is higher the larger the fraction of agents who support the rule of law. Then $\pi = \pi(x)$, where x denotes the share of agents who choose to strip assets, and π is a decreasing function of x .

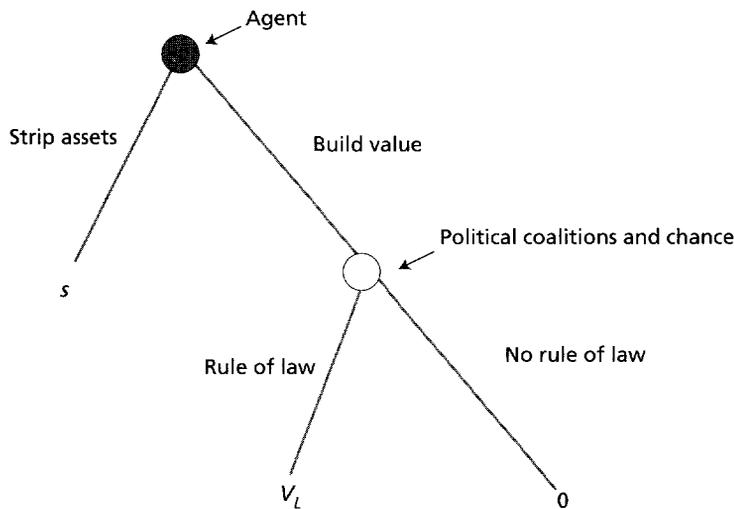
Time is divided into two periods. Asset strippers receive a payoff, denoted by s , in period 1 and nothing in period 2. Let F denote the distribution of the payoffs s across all agents.²⁰

The payoff to entrepreneurs comes in period 2. In present value terms it is V_L if the rule of law is in place and zero if it is not (figure 6). For simplicity, assume that V_L is constant and the same for all agents.

The expected payoff to building value is πV_L . Agents compare that value to their payoff from asset stripping. Thus, there is a threshold, denoted by s^* ,

$$s^* = \pi(x)V_L$$

Figure 6. Decision Tree of Agent with Ill-Defined Property Rights



Source: Author's illustration.

such that agents for whom $s > s^*$ choose to strip assets, and agents for whom $s < s^*$ choose to build value.

Any given value of the probability π maps into a unique value of the threshold s^* , and a unique fraction of agents, denoted x^* , for whom $s > s^*$:

$$x^* = 1 - F(s^*).$$

This model has elements of the entrepreneurship models discussed above. But now the spillovers are mediated *through the political process*. The higher the fraction of strippers, the smaller the support for the rule of law and therefore the higher the relative return to stripping. Thus, there can be multiple equilibria, as depicted in Figure 7.

Figure 7 is divided into two areas by the boldface curve, s^* . Below the curve, $s < s^*$ and so an agent with such a value of s is better off building value than stripping assets. Above the curve, the opposite is true. The curve $1 - x = F(s^*)$ is superimposed on the graph. These points are the payoffs to asset stripping of the $1 - x^{\text{th}}$ percentile of the population. The figure illustrates the special case where every individual would be better off as an entrepreneur under the rule of law than as an asset stripper. At $x = 0$, even the agent with the highest value of stripping assets, denoted \hat{s} , is better off building value than stripping assets: $\pi(0)V_L > \hat{s}$. Thus, if all controllers build value, the legal regime needed for investment is perceived as likely to emerge, $\pi(0)V_L$ is high, which makes $x = 0$ an equilibrium. However, at $x = 1$, even the agent with the lowest value of stripping assets, denoted \underline{s} , is better off stripping assets than building value. Thus, all controllers strip assets, the support for the rule of law is weak, $\pi(1)V_L$ is low, which makes $x = 1$ an equilibrium. In short, when few agents support the rule of law, the relative returns to building value *are* low and it *does* pay to strip assets.

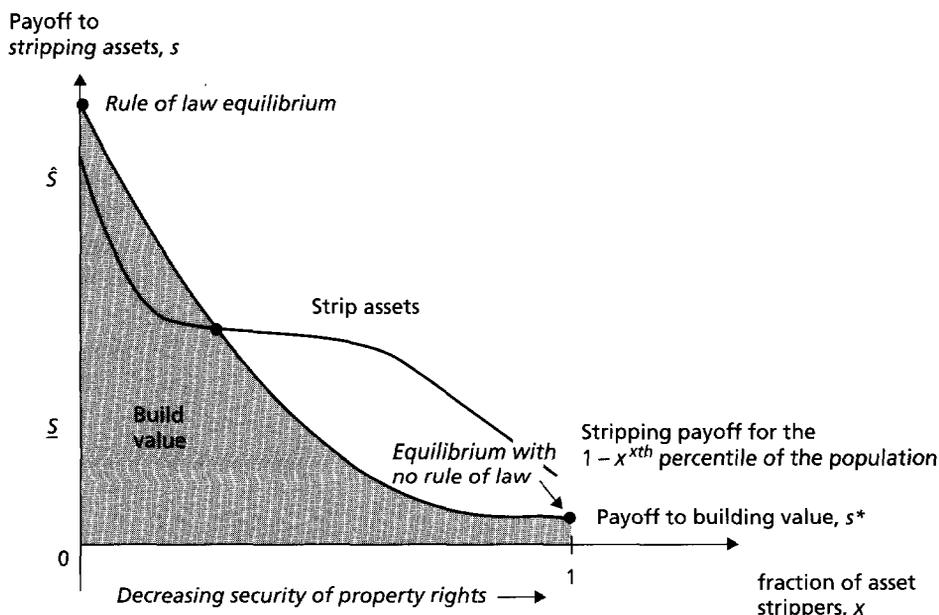
Some analysts have asked, “Why steal Gazprom (the giant Russian energy company) if you can make billions from it?” The model brings out the simple idea that if an agent’s current property rights to Gazprom are not respected in the future, then he cannot make billions from it (by normal business investments).

Conclusion

Developments in economics over the past 30 years have validated Rosenstein-Rodan’s basic intuition. A critical determinant of actions is one’s environment, one’s environment is endogenous, and an important aspect of the environment are the actions of other people. We used to believe that externalities would slightly distort the economy, but now we understand that the interaction of these slight distortions can produce large distortions. An economy may be in equilibrium in an environment with weak incentives for productivity, even though there is a better set of incentives for all individuals that would also be an equilibrium.

Modern theory has considerably extended Rosenstein-Rodan’s insight. We now see that some of the basic distinctions developed in neoclassical theory do not hold. There is no simple technology-based distinction between activities that produce externalities

Figure 7. Multiple Equilibria in Asset Stripping



Source: Author's illustration.

and those that do not, or between activities with public good properties and those with private good properties. The externalities that matter for welfare are not just direct interdependencies (Meade's beekeeper and apple farmer). There are many other classes of externalities with welfare consequences, an important example of which are those that arise in purely price-mediated interactions. Reflecting history, beliefs, and chance, certain behaviors are rewarded, others are not. Rewarded behaviors tend to increase relative to others, and that may further increase the rewards to those behaviors. Initial differences in circumstances or beliefs may not just persist, but be magnified over time.

In this view development policy is both easier and harder than before. It is easier because there is generally slack. A lack of resources need not be the fundamental constraint on development. A temporary policy intervention or seemingly minor shock may produce discontinuous change. But policy is harder because there is no single, easily identifiable source of failure waiting to be resolved.²¹ Development is not ensured by free markets and international trade, the transfer of capital, or the emergence of an entrepreneurial class. In some environments the entrepreneurial class becomes entrepreneurs to achieve predatory goals! Development requires complementary changes in the behavior of agents that not even the market can coordinate. Certain policies—policies that change beliefs, legal reforms, and the sequencing of policies—may contribute to coordinated changes to shift an economy to a better equilibrium.

Finally, there is an important difference between shallow interventions—which do not change the incentives of key players and so may fail to improve matters—and deep interventions—which entail irreversibilities. The sequencing of reforms

may make the difference between a deep intervention and a shallow one, as with privatization and anticorruption efforts.

Thus multiple equilibria, some with high incomes and others with low incomes, can result from strategic complementarities—an intellectually exciting idea with potentially enormous applications. Consider the simple adage, “If a man can . . . make a better mousetrap than his neighbor, though he builds his house in the woods, the world will make a beaten path to his door.”²² Can strategic complementarities matter even here? From the perspective of the literature reviewed in this article, the outcome predicted by the adage might still require that individuals expect that there will be enough users to make services in mousetrap repair affordable and search costs for those services low, that there be political support for property rights in mousetraps, that infrastructure sufficiently restrict the free movement of mice so that individuals have an incentive to buy their own mousetraps, and that there be a means for the disposal of dead mice.

The basic lesson from this literature is that externalities arise in many economic and social interactions and can give rise to coordination failures. Neither the market alone nor government alone can solve them. There are many misguided incentives in the private sector. There are different misguided incentives in the public sector. But it need not be that the misguided government cannot correct the misguided private sector. There may be a social equilibrium in which forces are balanced in a way that is Pareto improving relative to one in which the government’s hands are completely tied—and certainly better than one in which the private sector’s hands are completely tied.

Notes

1. This article uses the term *neoclassical theory* as short-hand for models that postulate rational agents who interact through a complete set of perfectly competitive markets. This narrow definition is for convenience only. There is no consensus on the scope of this term, which is sometimes used broadly to include any “systematic exploration of the implications of rational behavior in economics” (Laffont 1989, p. 6). In that broad sense all the models discussed in this article, except for one evolutionary model, are neoclassical, and information economics represents an intellectual revolution within neoclassical theory (Stiglitz 2001).

2. This can occur if there are agents for whom the incentive to choose an action is positively related to the total number of agents who choose that action—formally, there are *strategic complementarities*. Hoff and Stiglitz (2001a, appendix B) provide a simple mathematical taxonomy of models with *strategic complementarities*. Haltiwanger and Waldman (1991) show that a variety of phenomena in industrial organization and macroeconomics can be understood in terms of a framework where some agents are responsive to others’ actions, and some are not. Cooper (1999) surveys macroeconomic models with strategic complementarities.

3. Another channel of spillovers, outside the scope of this article, arises when one person’s preferences depend on another person’s preferences, as in *role model* effects. These effects can also give rise to coordination failures: the value placed on conformity can lead each individual to choose an action because others do. Yet because of intragroup interactions, another group with identical intrinsic characteristics can exhibit different behaviors that make all individuals better off. Durlauf (1999) illustrates this idea formally.

4. Besley and Coate (1998, pp. 151–52) provide a definition of political failure that is parallel to that of market failure. One begins in each case by defining technologically feasible utility allocations. For political failure, this reflects available policy instruments such as taxes,

transfers, and investments. Political institutions are then modeled. By analogy with market failure, a political failure occurs when equilibrium policy choices result in an outcome where it is technologically feasible (given available tax and transfer instruments, information, and so on) to implement a Pareto-improving policy, but that policy will not be an equilibrium choice.

5. In an economy with complete markets, pecuniary externalities correspond only to a movement from one Pareto efficient outcome to another. Such externalities give rise to changes in distribution that net out: gains by agents whose prices increase are exactly offset by losses to agents who must pay higher prices. There are no efficiency consequences from the allocation effects of price changes because equilibrium prices are always equated to marginal costs and benefits. But in an economy with incomplete markets, pecuniary externalities generally do not net out. A general framework is provided in Greenwald and Stiglitz (1986). One application of that idea (Hoff 1998) shows that in an economy where lenders cannot distinguish among borrowers who differ in their probability of default, pecuniary externalities from an improvement in technology can dissipate some, all, or more than all of the gains from the technological change. This occurs if the marginal borrower produces negative expected value (getting an implicit subsidy from the lower-risk, higher-quality borrowers). As the technological change induces the entry of marginal borrowers, the interest rate rises to reflect the lower average borrower quality, which hurts all borrowers.

6. Whereas most authors are associated with one strand or the other, Douglass North's work helped advance both. North's early work was in the Coasean tradition and pioneered its application to economic history. North's later work disparages prospects for understanding economic history as a more or less inevitable movement toward more efficient institutions: "Throughout most of history, the experience of the agents and the ideologies of the actors do not combine to lead to efficient outcomes" (North 1990, p. 96; see also North 1994). I owe to North the metaphor of the straitjacket (from his unwritten comments at a May 1999 World Bank conference, "Frontiers of Development Economics").

7. The literature on the relationship between wealth distribution and growth in economies with imperfect credit markets includes Galor and Zeira (1993), Banerjee and Newman (1993), Aghion and Bolton (1997), Piketty (1997), and Mookherjee and Ray (1998, 1999).

8. In a rejoinder to Coase (1974), van Zandt (1993) notes that even the early lighthouses in Britain did not operate in the "pristine 'private' world," but relied on government intervention to fix rates, grant monopolies, and collect user fees (p. 56).

9. Darwin (1993 [1859]) devotes two chapters to this idea (chs. 12 and 13).

10. An unstable equilibrium is one where the reaction curve is steeper than the 45 degree line.

11. For recent examples of evolutionary approaches to individual characteristics, see Bowles (1998), Fershtman and Weiss (1998), and Francois and Zabojsnik (2000).

12. Hoff and Sen (2000) survey the evidence.

13. We focus on the example of owning vs. renting, but other applications offer a choice between obtaining a formal ownership title vs. an informal one, or holding secure vs. insecure rights to property.

14. See, for example, Manski (2000), Azariadis and Drazen (1990), and Easterly and Levine (2000).

15. The key equilibrium condition equates the intertemporal marginal rate of substitution with the marginal product of household capital, which is a function of the initial endowment of household capital and the stock of capital in the county. The econometric analysis then tests the hypothesis that this function is decreasing in household capital and increasing in county capital.

16. Palfrey and Rosenthal (1984) present a model in which the larger is the number of potential beneficiaries of a discrete public good, the less likely the public good is to be supplied.

17. Amelina's comparison of two districts finds that this was more likely in the district where agriculture played a bigger role in the economy. In that district most politicians had risen through the ranks of the cooperative system. They had mastered the ability to transfer funds to and from the cooperatives and had a large political stake in meeting cooperatives' needs.

18. In practice a productive entrepreneur can always find ways of self-dealing. Even in developed economies there are well-known opportunities for tunneling (Johnson, La Porta, Lopez-de-Silanes, and Shleifer 2000). But there are tradeoffs between the two strategies: suppliers, minority shareholders, and employees will ultimately refuse to do business with a firm that defrauds them. Black, Kraakman, and Tarassova (2000) provide many examples of these tradeoffs faced by privatized Russian firms.

19. Black, Kraakman, and Tarassova (2000) provide extensive evidence for Russia of the asset strippers' flagrant use of the political process to undermine the rule of law.

20. There are many possible reasons for differences in s across agents who do not differ in terms of V_L . Asset stripping possibilities are greater in extractive sectors than in manufacturing. Among extractive industries, stripping possibilities are greater in firms where the capital investments needed for extraction are already in place. They are also greater where agents are able to collude with bureaucrats or to loot state banks, where the firms whose assets the agents control are highly leveraged, and where there is a ready market for the assets of the firm. So, agents who face the same rewards from entrepreneurship may face different rewards from asset stripping.

21. This despite a long history in development economics of arguing that there was such a source; see Adelman (2000).

22. Attributed to Ralph Waldo Emerson by Bartlett (1980). I consider the favorable case where people know about the invention.

References

- Acemoglu, Daron. 1995. "Reward Structures and the Allocation of Talent." *European Economic Review* 39 (January): 17–33.
- . 1997. "Training and Innovation in an Imperfect Labour Market." *Review of Economic Studies* 64 (220): 445–64.
- Adelman, Irma. 2000. "Fallacies in Development Theory and Their Implications for Policy." In Gerald Meier and Joseph E. Stiglitz, eds., *Frontiers of Development Economics: The Future in Perspective*. Oxford: Oxford University Press.
- Adsera, Alicia, and Debraj Ray. 1998. "History and Coordination Failure." *Journal of Economic Growth* 3 (3): 267–76.
- Aghion, Philippe, and Patrick Bolton. 1997. "A Theory of Trickle-Down Growth and Development." *Review of Economic Studies* 64 (2): 151–73.
- Amelina, Maria. Forthcoming. "Why Is the Russian Peasant Still a Kolkhoznik?" *Post-Soviet Geography and Economics*.
- Arnott, Richard, and Joseph E. Stiglitz. 1991. "Moral Hazard and Nonmarket Institutions: Dysfunctional Crowding Out or Peer Monitoring?" *American Economic Review* 81(1): 179–90.
- Azariadis, Costas, and Allen Drazen. 1990. "Threshold Externalities and Economic Development." *Quarterly Journal of Economics* 105 (2): 501–26.
- Banerjee, Abhijit J., and Andrew F. Newman. 1993. "Occupational Choice and the Process of Development." *Journal of Political Economy* 101 (2): 274–98.
- . 1998. "Information, the Dual Economy, and Development." *Review of Economic Studies* 65 (225): 631–53.
- Bardhan, Pranab, Samuel Bowles, and Herbert Gintis. 2000. "Wealth Inequality, Wealth Constraints and Economic Performance." In Anthony Atkinson and Francois Bourguignon, *Handbook of Income Distribution*. Amsterdam: Elsevier Science Ltd.
- Bartlett, John. 1980. *Familiar Quotations*. Boston: Little, Brown and Co.
- Basu, Kaushik. 1997. "On Misunderstanding Government: An Analysis of the Art of Policy Advice." *Economics and Politics* 9 (3): 231–50.

- . 2000. *The Social and Political Foundations of Economics: A Prelude to Political Economy*. Oxford: Oxford University Press.
- Basu, Kaushik, and Pham Hoang Van. 1998. "The Economics of Child Labor." *American Economic Review* 88 (3): 412–27.
- Becker, Gary. 1983. "A Theory of Competition among Pressure Groups for Political Influence." *Quarterly Journal of Economics* 98 (3, August): 371–400.
- Besley, Timothy, and Stephen Coate. 1998. "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis." *American Economic Review* 88 (1): 139–54.
- Black, Bernard, Reinier Kraakman, and Anna Tarassova. 2000. "Russian Privatization and Corporate Governance: What Went Wrong?" *Stanford Law Review* 51: 1730–1808.
- Bowles, Samuel. 1998. "Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions." *Journal of Economic Literature* 36 (1): 75–111.
- Boycko, Maxim, Andrei Shleifer, and Robert W. Vishny. 1996. "Second-best Economic Policy for a Divided Government." *European Economic Review* 40 (3-5): 767–74.
- Cheung, Steven S. 1973. "The Fable of the Bees: An Economic Investigation." In Henry G. Manne, ed., *The Economics of Legal Relationships: Readings in the Theory of Property Rights*. New York: West Publishing Company.
- Coase, Ronald H. 1937. "The Nature of the Firm." *Economica* 4 (November): 386–405.
- . 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3 (October): 1–44.
- . 1974. "The Lighthouse in Economics." *Journal of Law and Economics* 17 (2): 357–76.
- Cooper, Russell W. 1999. *Coordination Games: Complementarities and Macroeconomics*. Cambridge: Cambridge University Press.
- Cooter, Robert D. 1997. "The Rule of State Law and the Rule-of-Law State: Economic Analysis of the Legal Foundations of Development." *Annual World Bank Conference on Development Economics* 1996. Washington, D.C.: World Bank.
- . 2000. "Law from Order: Economic Development and the Jurisprudence of Social Norms." In Mancur Olson and Satu Kähkönen, eds., *A Not-So-Dismal Science*. Oxford: Oxford University Press.
- Darwin, Charles. 1993 [1859]. *The Origin of Species by Means of Natural Selection*. New York: Random House.
- de Meza, David, and J. R. Gould. 1992. "The Social Efficiency of Private Decisions to Enforce Property Rights." *Journal of Political Economy* 100 (3): 561–80.
- Dixit, Avinash K. 1996. *The Making of Economic Policy: A Transaction Cost Politics Perspective*. Cambridge, Mass.: MIT Press.
- Durlauf, Steven. 1999. "The Memberships Theory of Inequality: Ideas and Implications." In Elise Brezis and Peter Temin, eds., *Elites, Minorities and Economic Growth*. Amsterdam, Elsevier Science B.V.
- Easterly, William, and Ross Levine. 2000. "It's Not Factor Accumulation: Stylized Facts and Growth Models." World Bank, Development Research Group, Washington, D.C.
- Engerman, Stanley, and Kenneth Sokoloff. 1997. "Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies: A View from Economic Historians of the United States." In Stephen Haber, ed., *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800–1914*. Palo Alto, Calif.: Stanford University Press.
- Fershtman, Chaim, and Yoram Weiss. 1998. "Social Rewards, Externalities, and Stable Preferences." *Journal of Public Economics* 70 (1): 53–73.
- Francois, Patrick, and Jan Zabojsnik. 2000. "Why Doesn't Development Always Succeed? The Role of a Work Ethic." Working Paper. Queen's University, Kingston, Ontario.

- Galor, Oded, and Joseph Zeira. 1993. "Income Distribution and Macroeconomics." *Review of Economic Studies* 60 (1): 35–52.
- Greenwald, Bruce, and Joseph E. Stiglitz. 1986. "Externalities in Economics with Imperfect Information and Incomplete Markets." *Quarterly Journal of Economics* 100 (2): 229–64.
- Greif, Avner. 1994. "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies." *Journal of Political Economy* 102 (5): 912–50.
- Haltiwanger, John, and Michael Waldman. 1991. "Responders versus Non-Responders: A New Perspective on Heterogeneity." *Economic Journal* 101 (September): 1085–1102.
- Helpman, Elhanan, and Paul Krugman. 1985. *Market Structure and Foreign Trade*. Cambridge, Mass.: MIT Press.
- Hirschman, Albert O. 1958. *The Strategy of Economic Development*. New Haven, Conn.: Yale University Press.
- Hoff, Karla. 1994. "The Second Theorem of the Second Best." *Journal of Public Economics* 54 (June): 223–42.
- . 1996. "Market Failures and the Distribution of Wealth: A Perspective from the Economics of Information." *Politics and Society* 24 (4): 411–32.
- . 1998. "Adverse Selection and Institutional Adaptation." Working Paper 98-02. University of Maryland, Department of Economics, College Park, Md.
- Hoff, Karla, and Andrew Lyon. 1995. "Non-Leaky Buckets: Optimal Redistributive Taxation and Agency Costs." *Journal of Public Economics* 58 (3): 365–90.
- Hoff, Karla, and Arijit Sen. 2000. "Homeownership, Community Interactions, and Segregation." World Bank, Development Research Group, Washington, D.C.
- Hoff, Karla, and Joseph E. Stiglitz. 2001a. "Modern Economic Theory and Development." In Gerald M. Meier and Joseph E. Stiglitz, eds., *Frontiers of Development Economics: The Future in Perspective*. Oxford: Oxford University Press. 389–459.
- . 2001b. "Politics in the Aftermath of Privatization in the Transition Economies." World Bank, Development Economics Research Group, Washington, D.C.
- Isham, Jonathan, Deepa Narayan, and Lant Pritchett. 1995. "Does Participation Improve Performance? Establishing Causality with Subjective Data" *World Bank Economic Review* 9(2): 175–200.
- Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics* 108 (August): 577–98.
- Jalan, Jyotsna, and Martin Ravallion. 1998. "Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China." Discussion Paper 86. Boston University, Institute for Economic Development, Boston, Mass.
- Johnson, Simon, Dani Kaufmann, and Andrei Shleifer. 1998. "The Unofficial Economy in Transition." *Brookings Papers on Economic Activity* 2. Washington, D.C.: Brookings Institution.
- Johnson, Simon, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer. 2000. "Tunneling." *American Economic Review* 90 (2): 22–27.
- Kranton, Rachel E. 1996. "Reciprocal Exchange: A Self-Sustaining System." *American Economic Review* 86 (4): 830–51.
- Kremer, Michael, and Charles Morcom. 2000. "Elephants." *The American Economic Review* 90 (1): 212–34.
- Krugman, Paul. 1993. "Toward a Counter-Counterrevolution in Development Theory." In Lawrence H. Summers and Shekhar Shah, eds., *Proceedings of the Annual World Bank Conference on Development Economics 1992*. Washington, D.C.: World Bank.

- . 1995. *Development Geography and Economics*. Cambridge, Mass.: MIT Press.
- Laffont, Jean-Jacques. 1989. *The Economics of Uncertainty and Information* (translated by J.P. Bonin and H. Bonin). Cambridge, Mass.: MIT Press.
- Leibenstein, Harvey. 1957. *Economic Backwardness and Economic Growth*. New York: Wiley.
- Ljungqvist, Lars. 1993. "Economic Underdevelopment: The Case of a Missing Market for Human Capital." *Journal of Development Economics* 40 (April): 219–39.
- Manski, Charles. 2000. "Economic Analysis of Social Interactions." *Journal of Economic Perspectives* 14 (3): 115–36.
- Mason, Edward S. 1980. *The Economic and Social Modernization of the Republic of Korea*. Cambridge, Mass.: Harvard University Press.
- Matsuyama, Kiminori. 1995. "Economic Development as Coordination Problems." In M. Aoki, H-K. Kim, and M. Okuno-Fujiwara, eds., *The Role of Government in East Asian Economic Development*. Oxford: Clarendon Press.
- Miller, Marcus, and Joseph E. Stiglitz. 1999. "Bankruptcy Protection against Macroeconomic Shocks: The Case for a Super Chapter 11." World Bank, Washington, D.C.
- Mishan, E. J. 1971. "The Post-War Literature on Externalities: An Interpretive Essay." *Journal of Economic Literature* 9: 1–28.
- Mookherjee, Dilip, and Debraj Ray. 1998. "Persistent Inequality and Endogenous Investment Thresholds." Boston University, Department of Economics, Boston, Mass.
- . 1999. "Contractual Structure and Wealth Accumulation." Boston University, Department of Economics, Boston, Mass.
- Murphy, Kevin, Andrei Shleifer, and Robert Vishny. 1989. "Industrialization and the Big Push." *Journal of Political Economy* 97 (5): 1003–26.
- . 1993. "Why Is Rent-seeking So Costly to Growth?" *American Economic Review* 83 (2): 409–14.
- Myrdal, Gunnar. 1957. *Economic Theory and Under-developed Regions*. London: Duckworth.
- Newbery, David M. G., and Joseph E. Stiglitz. 1982. "The Choice of Techniques and the Optimality of Market Equilibrium with Rational Expectations." *Journal of Political Economy* 90 (2): 223–47.
- North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. New York: Cambridge University Press.
- . 1994. "Economic Performance through Time." *American Economic Review* 84 (3): 359–68.
- North, Douglass C., and Robert Paul Thomas. 1973. *The Rise of the Western World: A New Economic History*. New York: Cambridge University Press.
- Nurkse, Ragnar. 1953. *Problems of Capital Formation in Underdeveloped Countries*. Oxford: Oxford University Press.
- Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, Mass.: Harvard University Press.
- Palfrey, Thomas, and Howard Rosenthal. 1984. "Participation and the Provision of Discrete Public Goods: A Strategic Analysis." *Journal of Public Economics* 24 (2): 171–93.
- Piketty, Thomas. 1997. "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing." *Review of Economic Studies* 64 (2): 173–89.
- Radygin, A.D. 1999. "Ownership and Control of the Russian Industry." Paper presented at the "Conference on Corporate Governance in Russia," organized by the Organisation for Economic Co-operation and Development in cooperation with the U.S. Agency for International Development and the World Bank, Moscow, May 31–June 2.
- Ravallion, Martin, and Jyotsna Jalan. 1996. "Growth Divergence Due to Spatial Externalities." *Economics Letters* 53 (2): 227–32.

- . 1999. "China's Lagging Poor Areas." *American Economic Review* 89 (2): 301–05.
- Rodríguez-Clare, Andrés. 1996. "The Division of Labor and Economic Development." *Journal of Development Economics* 49 (April): 3–32.
- Rodrik, Dani. 1995. "Trade and Industrial Policy Reform." In Jere Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics*. vol. 3B. Amsterdam: North-Holland.
- . 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.
- Roland, Gerard, and Thierry Verdier. 1999. "Law Enforcement and Transition." Discussion Paper. CERGE-EI, Czech Republic.
- Romer, Paul. 1986. "Increasing Returns and Long-Run Growth." *Journal of Political Economy* 94 (October): 1002–37.
- Rosenstein-Rodan, Paul. 1943. "Problems of Industrialization of Eastern and Southeastern Europe." *Economic Journal* 53 (210–11): 202–11.
- Sah, Raaj Kumar, and Joseph E. Stiglitz. 1989. "Sources of Technological Divergence between Developed and Less Developed Countries." In Guillermo Calvo, Ronald Findlay, Pentti Kouri, and Jorge Braga de Macedo, eds., *Debt, Stabilizations and Development: Essays in Memory of Carlos Diaz-Alejandro*. Cambridge, Mass.: Basil Blackwell.
- Shleifer, Andrei, and Robert W. Vishny. 1998. *The Grabbing Hand: Government Pathologies and Their Cures*. Cambridge, Mass.: Harvard University Press.
- Sokoloff, Kenneth, and Stanley L. Engerman. 2000. "History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World." *Journal of Economic Perspectives* 14 (3): 217–32.
- Stiglitz, Joseph E. 1974. "Theories of Discrimination and Economic Policy." In G. von Furstenberg, Ann R. Horowitz, and Bennett Harrison, eds., *Patterns of Racial Discrimination*. Lexington, Mass.: D. C. Heath, Lexington Books.
- . 1998a. "The Private Uses of Public Interests: Incentives and Institutions." *Journal of Economic Perspectives* 12 (2): 3–22.
- . 1998b. "Reflections on the Theory and Practice of Reform." Paper presented to Stanford University, Department of Economics, Stanford, Ca.
- . 1993. "Comment on 'Toward a Counter-Counterrevolution in Development Theory.'" In Lawrence H. Summers and Shekhar Shah, eds., *Proceedings of the Annual World Bank Conference on Development Economics 1992*. Washington, D.C.: World Bank.
- . 2000. "Whither Reform? Ten Years of the Transition." In Boris Pleskovic and Joseph E. Stiglitz, eds., *Annual World Bank Conference on Development Economics*. Washington, D.C.: World Bank.
- . 2001. "The Contributions of the Economics of Information to Twentieth Century Economics." *Quarterly Journal of Economics* 115 (4): 1441–78.
- Tinbergen, Jan. 1967. *Economic Policy: Principles and Design*. Amsterdam: North-Holland.
- Tirole, Jean. 1996. "A Theory of Collective Reputations (with Applications to the Persistence of Corruption and to Firm Quality)." *Review of Economic Studies* 63 (1): 1–22.
- van Zandt, David E. 1993. "The Lessons of the Lighthouse: 'Government' or 'Private' Provision of Goods." *Journal of Legal Studies* 22 (1): 47–72.
- Zak, Paul J., and Stephen Knack. 1998. "Trust and Growth." Working paper 219. University of Maryland, Center for Institutional Reform and the Informal Sector, College Park, Md.

Comment on "Rethinking Aid," by Jan Willem Gunning, and "Beyond Rosenstein-Rodan: The Modern Theory of Coordination Problems in Development," by Karla Hoff

Gustav Ranis

These two professional and provocative articles are well off the beaten track of the development literature. Though I would have hoped for more breadth in dealing with changes in development theory and policy, both articles make interesting points.

Gunning's Article

Jan Willem Gunning's article addresses a main theme and a few side issues. The main theme is that Gunning accepts the conclusion of Burnside and Dollar (1997): that good policies are needed before aid can be effective. The question then is, how do donors and countries achieve good policies? Gunning maintains that it is not through ex ante conditionality, which has failed, but through ex post conditionality, or selectivity.

But why does ex ante conditionality not work? Gunning, Burnside, and Dollar seem to have been convinced by the reviews of the World Bank's internal evaluation units, as well as management surveys and academic assessments, that structural adjustment loans have a dismal record. Gunning cites four possible outcomes of conditionality and aid packages:

- Policy changes are not adopted.
- Policy changes are adopted but later reversed.
- Policy changes are adopted but would have been adopted anyway.
- Policy changes are adopted and stay in place.

According to Gunning, only the fourth outcome can be considered a success—and such cases have been few and far between.

A Few Problems

I have two problems with this assessment. The first is minor. Gunning considers the third outcome a failure because policy changes are not driven by conditionality. But

Gustav Ranis is professor of economics and director of the Center for International and Area Studies at Yale University.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

if a policy change is adopted—even if it would have been without a negotiated package—aid may still be crucial to buffer the possible pain arising from the change. Indeed, achieving success through policy dialogue, accompanied by “self-conditionality” (see below), seems ideal.

The second problem is more serious: I believe this premature burial of *ex ante* conditionality is fundamentally mistaken. First, I cannot resist pointing out that policy-based nonproject lending was not invented by the Bank but sometimes successfully deployed by the U.S. Agency for International Development in the 1960s in places like Chile, the Republic of Korea, Pakistan, and Taiwan (China), when the Bank still favored a “projects only” approach. But the larger point is that the frequent failure of structural adjustment conditionality has more to do with how it has been carried out than with its intrinsic merit. Its rejection reminds me of when people are unwilling to order a certain dish just because they have been to a bad restaurant.

The problem with *ex ante* conditionality as it is currently handled is that it lacks credibility. At the risk of caricaturing the process, *ex ante* conditionality packages have come to resemble a ritual dance where both parties are familiar with the music and practice well-rehearsed steps. With all the conditions in place and implied threats in case of noncompliance, both parties agree that the funds need to flow quickly—for the recipient because they are needed and for the donor because they are the main source of leverage and, moreover, are an indicator of banking “success.”

This approach to *ex ante* conditionality is hampered by:

- Inadequate nuancing of what needs to be done in a particular country, what few things can be done, in what sequence, and what risks should be insured against, over what time.
- The fact that it is still too centralized—for example, the Volcker Commission found that World Bank staff spent only 10 percent of their time in direct contact with least developed countries.
- Conditions and installments are usually overcome by the need to lend, a need deeply embedded in the flawed culture of the Bank—and even the International Monetary Fund, which hates to cut country programs. As a result aid may have the opposite effect of what is intended, as Devarajan, Dollar, and Holmgren’s (2000) study of structural adjustment in Africa demonstrates. As the pressure for reform is removed, this perverse impact of aid on decisionmaking becomes an extension of the Dutch disease. Others have called it an income effect, distinguishing it from the substitution effect of inducing the desired better policies.
- A lack of sufficient passivity on the part of the international community—which does not let recipient countries offer reform proposals based on domestic consensus and accompanied by self-conditionality agreements (to be differentiated from papered-over “ownership”).
- The unwillingness of international financial institutions to refuse to lend for several years if the conditions are not right, or to abandon loans midstream.

- The same donors' unwillingness or inability to adjust assistance in line with the risks, pain, and rewards of policy change and the need to ward off vested interests, including inside the government.
- The lack of attention paid to international institutions' potential role as knowledge banks, as emphasized by Gunning—allowing recipient countries to draw on them as they build domestic consensus on viable reforms. The provision of outside knowledge, as imperfect as it may always be, should be separate from the provision of dollars. Instead of denigrating the specialized agencies of the United Nations, they should be strengthened and accorded their proper role in the provision of know-how—when and if asked.
- Inadequate attention to the possible regionalization of *ex ante* conditionality. By this I mean increasing the involvement of regional development banks, which are closer to the local scene, and inducing neighboring countries (such as those in West Africa) to organize subregional self-conditionality following the post-World War II Marshall Plan in Europe.

Aid selectivity is Gunning's favored instrument. He sees policies as exogenous, so donors can reward "good" countries *ex post*. This notion troubles me for several reasons. Most important, learning occurs in repeated games, and policies inevitably become endogenous—so *ex post* becomes *ex ante* over time. Gunning seems to be aware of this, but overlooks it in his argument.

A second problem with selectivity is that it does not answer the question of what happens to poor people in badly performing countries. Donors cannot reach the poor through nongovernmental organizations or parachute drops if governments are opposed or not interested. So, donors need to do what they can even in less than ideal circumstances.

Third, apparently only Ghana, Mali, and Uganda would qualify for aid under Gunning's strictures. But Uganda's recovery has been quite recent. Was its marked change in policies really exogenous? And could it have been achieved without the *ex ante* provision of external resources?

Incidentally, I have no conceptual problem with an evaluation system based on outcomes rather than policy inputs, as suggested by Gunning. Indeed, I can see some advantage given that neither donors nor recipients ever grasp the complete truth. But we still must judge whether to reward countries when they have had some exogenous good luck, relative to others who tried to reform but were caught in an equally exogenous drought or terms of trade deterioration.

A Few Detours

Gunning also takes a few detours that are somewhat off the track of his main argument. For example, he unveils the special case where aid effectiveness depends not only on the quality of policies but also on structural vulnerability. I am puzzled: if adding such vulnerability variables as terms of trade fluctuations means that the aid and policy interaction variable is no longer significant, why the also? Are vulnerability and good policies really independent variables?

A second question concerns the claim that if there is a terms of trade decline, savings will be used to smooth consumption. If the permanent income hypothesis has not met empirical tests in industrial countries, it seems even less likely to apply to poor farmers in Sub-Saharan Africa. But perhaps the biggest puzzle is why, if Burnside and Dollar (1997) are right in the general case, Gunning finds it necessary to introduce this special case at all.

Later in the article Gunning discusses the high cost of taxation as a substitute for aid. I agree, though again I do not see a clear connection to his main theme, which focuses on policy change, not alternative sources of straightforward resource flows. Gunning advocates using unconditional aid for tax cuts or for transfer payments to favored elites. But what is the threshold where these elites exceed a critical share of the population? Moreover, I am not convinced that poor people are necessarily helped by across-the-board tax relief. They usually do not pay direct taxes; what they need is increased spending, especially on health and education.

Gunning presents another apparent obiter dictum. It is no surprise that Alesina and Dollar (1998) find that aid allocations are linked to former colonial ties. Donors—bilaterally or through their influence over multilateral agencies—often have objectives quite aside from achieving development goals. Regression analysis performed at the U.S. Agency for International Development in the 1960s found proximity to the communist bloc to be the most significant explanatory variable. In fact, country allocations are often determined by such political factors, while choices within countries are more likely to be guided by development goals.

A final comment on this stimulating and thought-provoking article: it is understandably somewhat Afrocentric. Would Gunning have expressed the same preference for ex post conditionality in South Asia or the Andean countries of Latin America?

Hoff's Article

Karla Hoff's article revisits Rosenstein-Rodan with the help of modern theory. Overall, it is a fascinating, erudite article—the author is clearly a superior theorist and craftsman. Still, I have a few concerns. For one thing, the article is geared toward looking for underdevelopment traps and downplays the analysis of a system's ability to escape from them. In this sense Rosenstein-Rodan, who discussed a possible escape through balanced growth—if not necessarily stimulated by market forces—was much more optimistic than Hoff. I agree that private rationality does not necessarily lead to social rationality, but neither are the two necessarily in conflict.

Turning to a more detailed assessment of Hoff's contribution, which has a distinctly Stiglitzian flavor, she argues convincingly against the inevitability of a neoclassical good equilibrium, and I agree. But neither is a bad equilibrium inevitable. Hoff presents a formidable list of possible coordination failures leading to traps:

- Markets are not perfect, and information asymmetries exist.
- Search costs are high.

- Even research and development can be subject to the prisoner's dilemma and a consequent race to the bottom.
- Returns to rents may exceed returns to productive activity—probably due to heavily intervened markets.
- *Institutional change does not necessarily lead to more perfect markets.*
- Even more heretically, market forces do not necessarily lead to a better equilibrium, given the high cost of enforcing property rights.
- An unequal distribution of wealth can lead to biases and distortions in factor markets.
- There may be a mismatch between investment in training and the requirements of a new technology.
- There may be predatory behavior on the part of government or a lack of social capital in civil society.

Unlike many of my colleagues, I do not have to be convinced that development is not an equilibrium enterprise. Even Kenneth Arrow readily acknowledges that the neoclassical good equilibrium, especially in a development context, is at best reached only over time. But Hoff promises to examine ways to counter coordination failures and harness spillovers with the help of policy reforms to enhance the chance for good equilibria. She refers to pitfalls and opportunities, but I found little reference to opportunities. Hoff seems to be trapped by the task of locating and explicating the meaning of a variety of traps.

In particular, I was surprised that Hoff did not refer to the new growth theory of Romer (1990) and Lucas (1988), which emphasizes the possibility of positive spillovers and externalities. While I am not impressed with the lack of empirical content in that theory, it does emphasize growth as an evolutionary process and focuses on opportunities. Admittedly, economic systems may be diverging instead of converging, but it is possible to achieve a good equilibrium and development success without convergence. Indeed, convergence makes unrealistic assumptions about all countries following the same production function and being subject to the same institutional constraints and knowledge frontier. That may hold only for typological neighbors, such as members of the European Union or states of the United States. Still, the theory should not be ignored.

Turning to Hoff's econometric test of rural China, I found somewhat doubtful the assumption of little factor mobility in response to different geographic opportunities, given the movement of workers to coastal areas and cities—especially now that migration prohibitions have been lifted. While I agree that county wealth affects household wealth through private income and public spending patterns, living standards and growth rate differences are more a function of different agricultural productivity and distance from ports, not wealth per capita (see Jalon and Ravallion 1998). Indeed, the richer Jalon and Ravallion model abandons county and household wealth and comes close to offering rather conventional multivariate analysis to explain consumption growth.

The basic question should presumably be what combination of markets and the state best reduces the risk of coordination failure—not what special assumptions are

needed to generate a trap. Deep interventions, not fully defined, are claimed to have permanent, irreversible effects, presumably leading to bad equilibria. Shallow interventions, not defined, presumably can be reversed and lead to positive opportunities. But this is not explored.

In conclusion, I accept that the behavior of other agents affects my welfare. But I am not convinced why this should most likely lead to a bad rather than a good equilibrium, counter to what new growth theory tells us. Hoff is clearly a capable professional. I just hope that in the future she does not take her "dismal science" union card as seriously as she has in this article.

References

- Alesina, Alberto, and David Dollar. Forthcoming. "Who Gives Foreign Aid to Whom and Why?" *Journal of Economic Growth*.
- Burnside, Craig, and David Dollar. 1997. "Aid, Policies and Growth." Policy Research Working Paper 1777. World Bank, Washington, D.C.
- Devarajan, Shanta, David Dollar, and Torgny Holmgren. 2000. "Aid and Reform in Africa." World Bank, Development Research Group, Washington, D.C.
- Jalan, Jyotsna, and Martin Ravallion. 1998. "Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China." Discussion Paper 86. Boston University, Institute for Economic Development, Boston, Mass.
- Lucas, R. E. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (July): 3–42.
- Romer, P. M. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98 (October): S71–S102.

Comment on “Rethinking Aid,” by Jan Willem Gunning, and “Beyond Rosenstein-Rodan: The Modern Theory of Coordination Problems in Development,” by Karla Hoff

Charles C. Soludo

Especially for the 48 least developed countries, few topics are more important than those raised by Jan Willem Gunning and Karla Hoff. Despite more than 50 years of development economics and official development assistance—along with 20 years of special attention to Sub-Saharan Africa—the track record is an embarrassment. Since the early 1970s only diamond-rich Botswana has moved out of the group of least developed countries, this despite the fact that official development assistance to these countries has averaged 19 percent of GDP a year.

In fact, during this period more countries became least developed countries than graduated from the class. The average person in these countries was richer in 1979 than in 1999. Severe poverty continues to deepen. And the debt burden from earlier aid circumscribes prospects for sustainable development as well as the effectiveness of current and future aid. For these countries it is as if the insights from development economics and the role of development finance have been incomplete or entirely wrong. Thus the entire approach to development must be reconsidered, and Hoff and Gunning should be congratulated for the job they have done dealing with voluminous, difficult subjects.

But while Hoff and Gunning generally do a good job of addressing underdevelopment traps, their analyses suffer from two defects. First, the analyses center on the nation-state—so there is no recognition of the implications of globalization. The nation-state is grossly inadequate for understanding why underdevelopment traps have become localized in certain regions and why aid is often ineffective. Second, the authors’ policy conclusions are somewhat simplistic and couched in an extremist either-or framework. As a result they overlook complementarities in the causes of and solutions to underdevelopment traps and ineffective aid.

Gunning’s Article

I have several concerns about Gunning’s article. First, its title advertises more than the article delivers. Rather than addressing the larger issue of aid, it focuses on the

Charles C. Soludo is professor of economics at the University of Nigeria.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

functions of aid. It reviews the recent debate on the role of aid in providing finance, changing policies, transmitting knowledge, and providing insurance against shocks. Needless to say, some of the conclusions are controversial.

But Gunning omits several important issues in the debate on the delivery and effectiveness of aid. Among these are basic issues about the domain of aid—that is, the amounts to be delivered by communities, the central government, and regional or global actors. (Put another way, this is the debate on national and regional or international public goods.) Gunning also misses an important debate on donor coordination (see Kanbur, Sandler, and Morrison 1999). In addition, the article skates around the question of what aid should finance beyond tax relief and a knowledge bank. Finally, the relationship between debt and current and future aid effectiveness is overlooked. In countries where debt payments exceed government spending on health and education, or where debt rescheduling eats up more than half the time of senior civil servants, it is hard to evaluate aid without addressing these issues.

Second, Gunning argues that *ex ante* conditionality does not work, and should be replaced by *ex post* conditionality, or selectivity. But this prescription is based on the shaky premise of an exogenous policy regime. In most of the least developed countries, policy choices—and more important, the sustainability of reforms—are not determined by exogenous political will. Policies are circumscribed by a number of other issues, including debt, expected aid, and exogenous economic shocks. While aid cannot buy sustained reform from unwilling governments, willing governments of least developed countries cannot sustain reform without aid. It is a chicken and egg problem.

For example, frequent reversals of trade liberalization have little to do with political will for reform and much more to do with the macroeconomic incompatibility of liberalization. None of Africa's least developed countries has been able to sustain trade liberalization without *ex ante* aid or promises of aid to smooth the fiscal and balance of payments problems that liberalization entails. So, if donors are looking for a proven track record with trade liberalization before delivering aid, they are unlikely to find one in Africa. Ultimately, aid would continue to be a gamble—with an *ex ante* commitment being the relevant variable. Furthermore, *ex ante* conditionality has not failed in all cases, and the swing to selectivity is a naive response to an imperfect delivery process.

In my view selectivity is donors' response to declining aid. The argument seems to be that if there is not enough aid to go around, why not just give it to countries that can make the best use of it? But selectivity does not solve the problem of declining aid. If all the least developed countries were to implement the policies required by selectivity, current aid would be too thin to be effective.

A third issue involves the signaling role of aid. Gunning argues that donors could use selectivity to signal that aid is tied to success. But such signaling does not mesh with country incentives, because success results in higher debt payments. Furthermore, the argument for using aid to lock in policies in recipient countries presumes certainty on what constitutes the best policy options. Because donor pref-

erences are subject to fads and fashions, efforts to lock in “good” policies run the risk of maladjusting the least developed countries—as the past 20 years have shown. Gunning acknowledges this uncertainty on what works best. Put differently, the signaling role of aid would not be all that different from ex ante conditionality because it would signal which behaviors are rewarded.

My last comment on Gunning pertains to his argument that aid be used for tax relief. True, there is a strong argument for using aid to provide temporary tax relief at early stages of growth. But it cannot be generalized. Aid and taxation should be seen as complements, not substitutes. Domestic resource mobilization does not always imply higher tax rates. It could involve higher compliance ratios—or lower rates with a broadened base. For example, lower tariffs could reduce the incentive for smuggling, raising government revenue. Resource mobilization also depends on initial conditions. In a country where tax revenue accounts for 5–10 percent of GDP, it is possible to raise revenue without distorting production.

Hoff's Article

Hoff summarizes the evolution of modern theories of development, especially the broader understanding of sources of spillovers. She emphasizes developments in institutional economics, path dependence, and views on externalities and public goods. These models show that multiple equilibria can occur—some of them desirable and some of them underdevelopment traps—and that the market may fail to deliver a good equilibrium. Thus coordination failures are important.

Hoff evaluates new insights on the nature of the traps, which in the 1950s necessitated official development assistance, and concludes that the new view dictates not more resources but less resources and deeper interventions (requiring more skills). This conclusion requires further elaboration: it is not clear from the article what less resources would entail. More precisely, couching the issue in terms of either resources or legal changes and the like (deep intervention) can send the wrong message. From the analysis, both are required: they are complements, not substitutes. Hoff acknowledges this complementarity in several places. For example, she argues that development requires complementary changes in the behavior of agents that not even the market can coordinate.

Hoff goes on to recommend temporary legal reforms, policies that change beliefs, and changes in the distribution of wealth, which may contribute to coordinated changes to shift an economy to a better equilibrium. But the effectiveness of such efforts may be circumscribed by the same factors that Hoff refutes as being fundamental to development—free markets and international trade, the transfer of capital, and the emergence of an entrepreneurial class. Denying the importance of these variables does more to obfuscate than clarify the issues. For example, such a position ignores the causal two-way links between beliefs, culture, and the production structure on one hand, and the alteration of the production structure and the variables described as not being fundamental to development on the other. Frankly, the tragedy of the development discourse of the past 50 years is that every little addition has been

dramatized as the ultimate insight instead of part of the accumulated body of knowledge. Complementarity is the key word that many analysts fail to internalize.

My second point is that Hoff takes a narrow view of underdevelopment traps, largely confining her analysis to nation-states. If globalization is taken into account, and traps are analyzed within the context of the global village, both the breadth and depth of the traps would change. The analysis of multiple equilibria would apply as much to a nation-state, a region of a country, or the world. Policy responses—shallow or deep interventions, or combinations of skills and resources—would differ depending on the framework of analysis. If Rosenstein-Rodan framed his analysis in the context of a closed economy, Hoff's is to a large extent guilty of a similar framework.

An open economy framework is crucial for understanding traps and their remedies because of the asymmetries of globalization. Globalization enables productive resources to move freely while populations are confined to certain locations. Given the poorer initial conditions of certain locations and path dependence, destitution seems to be localized. Moreover, the cost of breaking out of a vicious circle seems to be getting higher by the day. For example, countries that fall short of the capital required for self-sustaining growth are, given path dependence, condemned to low-level equilibrium. As a result investments quickly move from poorer to richer countries. Thus the Biblical injunction is fulfilled that "to he who has, more shall be given and to he who has not, even the little he has shall be taken away and given to he who has more."

This process explains the increasing divergence between least developed and OECD countries over the past 40 years. And despite the liberalization of current accounts and capital mobility, returns are not equalized. Nor does arbitrage occur to equalize returns: there is still a home bias for capital or even the "ethnicity" of capital. The United Nations Conference on Trade and Development estimates that investment returns are highest in Africa, yet foreign direct and portfolio investments have not flowed to the region. Without conscious efforts to break out of the traps, poor economies might remain mired in perpetual poverty.

Globalization Tax and Trade as Remedies?

From the above discussion it is evident that low-level equilibrium traps are as much a feature of national economies as of the global system. Getting out of the traps requires good policies. Poor countries need stronger institutions, deeper reforms, and better sequencing. But they also need increased global coordination and a significant big push—given the drains on the system due to poverty, much larger resource transfers are needed. This same insight underscored the development economics of the 1950s. But the main reason past resource transfers failed was that their size and timing were determined more by the whims of the donors than by the development needs of the recipients. In addition, there is little international coordination to make aid work for development and help poor economies move from aid dependence to sustainable growth.

Aid flows are falling. Despite growing needs and better policies in poor countries, aid dropped 23 percent in real terms in the 1990s. Aid is falling because its political and commercial motivations have changed. With the World Trade Organization, there is little incentive to “buy” trade with aid. With the end of the Cold War, there is little incentive to dole out aid for strategic interests. And countries hardly provide charity with aid, except occasional humanitarian responses.

Because low-level equilibrium traps will persist without significant aid injections, the challenge is ensuring increased resource flows and sustaining a move out of poverty. Because the current system based on altruism does not work, a global mandatory transfer mechanism is needed. An example of such a mechanism is the proposal for a globalization tax (Soludo 1999).

The case for a mandatory transfer is intuitive. Multilateral arrangements—the World Trade Organization, the Bretton Woods institutions, the United Nations system—already harmonize policies and institutions, but without a governance structure to ensure mandatory transfers to pull along left-out countries. The current global system is akin to a federal state without free movement of people, and with no mandatory transfers to poorer regions.

This system needs serious rethinking to develop a structure for implementing the transfers. For example, further attention could be given to the suggestion by the U.S. Congressional committee on the Bretton Woods institutions that the World Bank become solely a grantmaking institution and the International Monetary Fund serve as the global lender of last resort (world central bank). A globalization tax, perhaps set as a percentage of countries’ GDP or realized as proceeds from a global Tobin tax, could be paid to the World Bank. In turn, grants would be made to the poorest countries based on agreed criteria, including the level of underdevelopment, commitment to reforms, and need for regional and international public goods.

Beyond increased mandatory transfers to poor regions, the multilateral system needs better coordination. Even with increased mandatory transfers, globalization may not lead to good equilibria at the global level. Better coordination requires action at two levels. On the one hand there is a need for increased capacity and capability building—especially to diversify production and build export competence. Here aid can help provide skills and act as a knowledge bank (as emphasized by Gunning).

But different coordination is also required, and entails providing the poorest countries with preferential trade opportunities. Unless small, poor economies have unlimited access to OECD markets it is difficult to see how they can achieve sustainable growth. Since populations will not move freely across borders, trade preferences can be a powerful stimulus for global income distribution. OECD countries provide agricultural subsidies equal to Africa’s GDP, and this agricultural output is dumped in Africa. Though Africa is supposed to have a comparative advantage in primary commodities, it cannot sustain a competitive advantage in the face of such dumping. Furthermore, even the little that Africa produces faces enormous obstacles because of nontariff barriers imposed on such goods by OECD countries. Thus, to escape the low-level equilibrium and empower African producers to reap appropriate returns from their labor, unrestricted access to Western markets is crucial.

Without mandatory transfers to the poorest countries and coordination that gives producers in these countries unconditional access to Western markets, it is hard to see how underdevelopment traps can be overcome or future aid made more effective. If these ideas sound too idealistic, so be it: nearly every idea that later became orthodoxy started that way. The problem is not so much the soundness of the idea but whether rich countries have the political will and courage to think and do things differently.

References

- Kanbur, Ravi, Todd Sandler, and K.M. Morrison. 1999. "The Future of Development Assistance: Common Pools and International Public Goods." Policy Essay 25. Overseas Development Council, Washington, D.C.
- Soludo, Charles C. 1999. "Towards an Evaluation of the Proposals for the New Global Financial Architecture." Paper presented at Oxfam America workshop on the New Global Financial Architecture, Annapolis, Md., May.

*Crises and
Recovery*

Shaken and Stirred: Explaining Growth Volatility

William Easterly, Roumeen Islam, and Joseph E. Stiglitz

This article explores the relationship between volatility in economic growth and various institutional factors. Its main hypothesis is that in explaining volatility, the traditional macroeconomic literature has overemphasized such factors as wage and price rigidities and underemphasized factors relating to the financial system. The financial system may act as a stabilizer that helps to cushion consumers and producers from the effects of economic shocks, or it may magnify these effects and thus increase growth volatility. Some of these destabilizing effects operate through the cash flow and balance sheets of banks and firms, leading to increased credit rationing. An empirical investigation reveals that wage and price rigidities are not important in explaining growth volatility. It also shows that the financial system generally acts as a stabilizer and reduces growth volatility. But the relationship is nonlinear. As the financial system grows, its risk-enhancing characteristics can result in higher growth volatility.

The world's economic history is replete with recessions and depressions. Crises have been a constant of market capitalism—from the bursting of the British South Sea bubble and the French Mississippi bubble in 1720 (which at least one economic historian claims delayed the industrial revolution by 50 years), to the depressions of the 1870s and 1930s in the industrial economies, to the debt crises of middle-income Latin American countries and low-income African countries in the 1980s, the collapse of output in the formerly socialist economies in the 1990s, and the East Asian financial crisis in 1997–98.¹ Add to these the collapses that have accompanied non-economic shocks—wars, fires, pests, floods, droughts, earth-

William Easterly is senior adviser in the Development Economics Research Group at the World Bank. Roumeen Islam is Director of the World Development Report team in the Development Economics Group at the World Bank. Joseph E. Stiglitz is professor of economics at Stanford University. Some of the ideas in this article were presented as the Michael Bruno Memorial Lecture at the 12th World Congress of the International Economics Association in August 1999. The authors are grateful for comments by participants in that event. They are also grateful to Mattias K. Lundberg and Alexander Lebedinsky for their valuable research assistance.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

quakes, hurricanes, volcanic eruptions—and it is a wonder that there is economic security anywhere.

Recent economic crises have often gone hand in hand with financial crises, which have become increasingly frequent and severe in developing countries over the past quarter century. These financial crises have had different causes and natures. For example, those characterizing the debt crises of the 1980s were precipitated by profligate governments with large cash deficits and uncontrolled monetary policies. The more recent ones occurred in countries that for the most part were following prudent macroeconomic policies and some of which had quite sophisticated institutional arrangements.

The marked differences between the downturns in Latin America in the early 1980s and those in East Asia in the late 1990s (and the Mexican crisis of 1994–95) mean that we need a general framework for thinking about macroeconomic fluctuations—one that can encompass differences among countries. There are ample reasons for trying to better understand the determinants of economic volatility: volatility is important not just because of its short-run adverse effects on the poor, but also because of its negative correlation with economic growth.

We attempt to set forth a framework for thinking about growth volatility that is general enough to incorporate the important structural, institutional, and policy variations among countries that might account for differences in their macroeconomic performance. We focus particularly on the role of the financial sector. We first discuss the importance of short-run dynamic effects in determining long-run outcomes, and the role of the financial sector, elements that have not been sufficiently incorporated into traditional macroeconomic analysis. We then look at the data, which reveal interesting aspects about the determinants of volatility—most notably, the importance of the financial sector.

Dynamics, Financial Variables, and the Standard Competitive Model

The starting point of modern macroeconomics is the competitive equilibrium model, in which all resources are not only fully employed, but deployed efficiently. Fluctuations in output therefore reflect changes in inputs (say, the desire of workers to work) or changes in technology—the relationship between inputs and output. While these real business cycle theories provide plausible explanations for variability in the rate of growth, they are less successful in providing persuasive explanations of economic downturns in a large, closed economy such as the United States. Can one really believe that these factors caused the Great Depression, or even the Reagan recession? That the reduced employment reflected a sudden desire of workers to enjoy more leisure, a desire that quickly changed again a couple of years later? For small, open economies adverse terms of trade shocks can have much the same effect as negative technology shocks, and this is one of the important differences between the macroeconomics in these economies and that underlying some of the traditional closed economy models.

Employment and output fluctuations inevitably relate to shocks and to the way the economy copes with those shocks. They are determined by the extent to

which the individually rational actions of firms and households and the policy interventions of governments add up to collective behavior that either brings the economy quickly back to full employment and efficient resource utilization or does not. These issues are particularly complicated because what is viewed to be individually rational on the part of firms and households depends on their beliefs about both the behavior of each other and the policy regime of the government. The policy regime in turn may depend on the government's beliefs about the behavior of firms and households. And the shocks themselves are at least to some extent endogenous, determined by outsiders' beliefs about the economic structures. Thus modern macroeconomics is concerned with the dynamics of quite complex systems.

Classical business cycle theory provides a different perspective: it sees the economy as described by a set of difference or differential equations, which exhibits cyclical behavior. The most famous examples of such equation sets are Paul Samuelson's multiplier accelerator model and J. R. Hicks's business cycle theory. The fundamental objection to these mechanistic approaches—beyond the unpersuasiveness of some of the underlying technological assumptions (such as the accelerator)—is that if they were true, downturns would be predictable. Governments could, through monetary and fiscal policy, take countervailing measures.

For nearly half a century after World War II attention centered on the downward rigidity in money wages and prices as a possible explanation of economic fluctuations. Rigid real wages provided an easy explanation of unemployment—a leftward shift in the demand curve for labor immediately turned into unemployment. And the leftward shift in the demand for labor could be explained by the falling demand for goods, itself explained by rigidities in intertemporal prices—for example, rigidities in the interest rate, which monetary policy seemingly could not bring down or could not bring down enough to stimulate consumption and investment.

Subsequent work has focused on amplifying the reasons for nominal and real wage rigidities (menu costs, efficiency wage theory, portfolio theories of adjustment) and on finding deeper explanations beyond the liquidity trap for the failure of monetary policy to bring down interest rates (for example, risk-averse behavior of banks, especially when confronted with excessively tight regulatory oversight).

Dynamics

Even within that traditional framework, however, much of the standard analysis has failed to emphasize some important first-order effects, such as the dynamic consequences of wages and prices falling. These may result in short-run adverse effects that appear earlier and are more dominant than the comparative static effects, which have been the primary focus of attention. The difference is not just a matter of exposition: the dynamics of adjustment may have an effect opposite that predicted by a comparative static analysis.

For example, it is usually asserted that a fall in prices will raise consumption through the real balance effect. The more precise statement is, "lower prices would

be associated with higher consumption.” But there is typically no instantaneous jump in consumption. Falling prices mean that at any level of nominal interest rates, real interest rates rise—and presumably investment falls (overall demand may also fall). Similarly, it is often stated that lower wages may be associated with higher employment. But to go from one level of wages to a lower level, wages need to be falling. If falling wages lead workers to reduce consumption, the net effect on aggregate demand and employment could even be negative.

Some strands of research (recent as well as predating Keynes) have focused on differences in adjustment speeds (see, for example, Stiglitz 1999) as well as on distributive effects that arise from price changes, especially those against which individuals cannot be insured (reflecting incomplete contracts). We are increasingly aware that income effects arising from distributional changes can often overwhelm substitution effects arising from price changes. This is especially true when there are asymmetries in the adjustment of real variables. For example, it is easier, less risky, or less costly to contract than to expand the use of some inputs.²

The Importance of Financial Institutions

There has been growing recognition that wage and price rigidities may not be the only, or even the most important, departure from the standard competitive equilibrium model relevant for explaining economic fluctuations. Models based on wage and price rigidities become unpersuasive if countries have both flexible wages and flexible prices and still exhibit high volatility in growth. We need to ask whether this high volatility can be explained simply by the fact that the countries are exposed to more shocks (or have a less diversified economy) than others or whether it is explained by other aspects of their structure or policy regimes.

This question leads us to a second difference between the new perspective and traditional macroeconomic analyses. In the traditional analyses institutions (other than labor market institutions that give rise to wage rigidities) play no role. But one of our central theses is that earlier studies have not paid sufficient attention to dynamic effects arising from the wealth and cash flow constraints of firms and financial institutions, such as banks and securities markets. (Under neoclassical theory, these constraints simply do not exist.) Financial institutions have profound effects on the behavior of firms (on how they cope with shocks, for example), and firms have profound effects on the behavior of financial institutions.

FIRM WEALTH EFFECTS. When negative net worth shocks are large enough (such as when there are interest rate shocks), firms may go into distress—that is, they may be in bankruptcy or on the verge of it.³ Because of the complex credit relationships among firms—most supply credit to customers, suppliers, or both—the bankruptcy of one firm can set off a “bankruptcy chain,” weakening other firms that depend on it and possibly pushing some into bankruptcy. Thus the likelihood of bankruptcy becomes a systemic concern (see Orszag and Stiglitz 1999), and negative effects on output and growth may materialize.

As more firms go into distress, the number of nonperforming loans increases and the financial positions of financial institutions deteriorate.⁴ Theory and evidence both support the hypothesis that firms act in a risk-averse manner and that the effective degree of risk aversion is affected by their wealth—for example, how close a firm is to bankruptcy.⁵ Adverse net worth shocks to financial institutions reduce their ability and willingness to bear risk—that is, the shocks reduce the amount they are willing to lend at any interest rate. Certain groups of borrowers may be excluded from the market as a result, an outcome that may exacerbate the economic downturn.

CASH FLOW CONSTRAINTS. In standard economic theory cash flow (or liquidity) constraints simply do not exist: anyone with good future prospects can get access to funds. But there is evidence, especially for small firms, that cash flows do have large effects on firms' decisions, for example, those on investment and, in extreme cases, even production. Imperfections in the equity market (both adverse selection and incentive concerns) lead to what might be thought of as equity rationing. At the very least, the costs of issuing new equity may be very high, making firms reluctant to engage in this form of finance even when they cannot obtain loans (see Myers and Majluf 1984; Greenwald, Stiglitz, and Weiss 1984; and Hellmann and Stiglitz 2000). Moreover, under equity rationing firms will be unable to diversify their risk well and will thus be more risk averse.

The size of this “financial sector” effect is determined in large part by the extent of the economy's integration into global capital markets. Weaknesses in the country's own financial institutions may matter little if firms have easy access to banks abroad. But while a high degree of capital account openness could in principle smooth a country's adjustment to a shock, it might also expose the country to another, adverse source of dynamic reaction. Investors observing the weakening condition of firms and financial institutions in response to the shock might decide to pull their (short-term) money out of the country, further weakening both firms and financial institutions (for example, by further weakening the currency) and possibly inducing a crisis.

A negative shock to the capital account will have adverse effects on the terms at which firms can get access to funds (which will adversely affect both liquidity and net worth), effects that may be exacerbated by the presence of credit rationing. The increased uncertainty about firms' balance sheets caused by the economic disturbance may lead to greater credit rationing and to further contractions in demand (investment, including inventories) as firms attempt to increase their liquidity.

By underemphasizing the financial sector, particularly the dynamics within the sector, the standard stories leave out much of the richness of the macroeconomic adjustment process and perhaps much that is of first-order importance. Many of the seeming anomalies (deviations from model predictions) that we observe in the real world can be explained by a model that incorporates a variety of the effects discussed here. For example, consider the seeming anomaly of fluctuations in output in small, open economies for which aggregate demand should not be a central problem (as long as the exchange rate is reasonable). We can explain this phenomenon by focusing on such factors as interruptions in the flow of credit and high interest

rates, which can combine to force many firms into bankruptcy, shifting the market supply curve to the left.

Endogeneity of Institutions and Shocks

Views that recognize the importance of institutions and the dynamics also emphasize the endogeneity of many factors previously taken to be exogenous—including institutions and shocks. Thus countries such as those in East Asia or Western Europe may have more financial depth in part because they experience fewer shocks. Had these countries faced the shocks experienced elsewhere, firms would have been unwilling to undertake the risks associated with high debt strategies, households would have been unwilling to save in financial assets, and governments might have been unwilling to provide the implicit or explicit insurance that made those risks more bearable. But countries in which firms have high debt-to-equity ratios and financial institutions are highly leveraged may “invite” shocks—that is, they may be highly susceptible to changes in the perception of their economic future.

But clearly not everything can be endogenous—or at least cannot be perceived that way by policy economists. Governments can be thought of as adopting a policy regime, such as whether and when to open the capital account or liberalize trade (though from the perspective of political economy, even the policy regime can be thought of as endogenous). Governments can decide whether to deregulate financial institutions. They may be able to decide—within constraints—on the macroeconomic regime. Decisions about these policy regimes should be sensitive to the characteristics of the economy—and the subtleties of dynamics. Certain forms of liberalization may promote economic growth and stability under certain circumstances, while similar policies pursued under other circumstances may slow growth and contribute to instability. More flexible wages and prices may increase economic stability under certain circumstances, while under other circumstances moves to enhance wage flexibility could exacerbate an economic downturn. On average, greater openness may be good, but in particular circumstances it may increase volatility.

To ascertain which of the effects discussed are more important, and how their importance compares with that of the factors traditionally emphasized—wage and price rigidities—we must turn to the data. The data cover both OECD and developing countries from 1960 to 1990.

What Do the Data Show?

Mean growth in developing countries is lower than that in OECD economies, and much more volatile (table 1). These two findings are consistent with the empirical studies showing that the partial correlation between growth and the volatility of growth is negative (for example, Ramey and Ramey 1995).⁶ Employment also is much more volatile in developing than in developed economies.

Table 1. Real Output Growth and Volatility in Real Growth and Employment

Variable	Developing countries		High-income OECD countries			
	Mean	Number of observations	Mean	Number of observations	t-statistic for difference in means	P-value
Growth	0.007	163	0.027	23	-5.659	0.000
Standard deviation of growth	0.061	163	0.026	23	9.779	0.000
(Median standard deviation of growth)	0.052		0.022			
Standard deviation of employment	0.098	83	0.035	21	6.652	0.000

Source: Authors' calculations.

Table 2. Some Empirical Differences between Developing and High-Income OECD Countries

Variable	Developing countries		High-income OECD countries			
	Mean	Number of observations	Mean	Number of observations	t-statistic for difference in means	P-value
Standard deviation of real wage index	2.119	90	1.883	21	0.833	0.410
Standard deviation of real wage changes	1.197	85	0.321	21	8.116	0.000
Standard deviation of fiscal balance	3.916	111	2.438	23	3.978	0.000
Credit to private sector/GDP	25.280	148	64.023	22	-6.441	0.000
Standard deviation of credit to private sector/GDP	9.179	148	21.206	22	-5.101	0.000
M3/GDP	38.065	148	65.805	21	-4.766	0.000
Standard deviation of M3/GDP	10.572	148	12.320	21	-0.785	0.440
(Imports/exports)/GDP	79.285	154	60.972	24	2.399	0.022
Standard deviation of inflation	0.219	148	0.043	23	6.234	0.000
Private capital flows/GDP	1.722	146	0.372	22	2.743	0.009
Standard deviation of private capital flows/GDP	2.662	138	2.311	22	0.808	0.420
Standard deviation of terms of trade changes	0.123	117	0.041	23	9.688	0.000
Standard deviation of money growth	0.219	148	0.077	20	6.757	0.000

Source: Authors' calculations.

There is a paradox, however: developing countries have greater volatility in real wages than do OECD economies (table 2). In the comparison of these two groups of countries there is little to support an explanation of employment and output fluctuations based purely on nominal wage rigidities—developing countries have more flexible real wages, yet they also have greater volatility in output and employment.

This may suggest that the demand effects of real wage changes dominate the supply effects, or that there is reverse causation from output volatility to real wage volatility (we address this issue in the regressions).

What then could explain the greater growth volatility in developing countries? There is little empirical or theoretical work on what determines volatility in growth rates.⁷ Theoretical considerations suggest that greater openness to trade might expose a country to more external shocks while leaving it less vulnerable to internally generated shocks. Greater openness of the capital account might in principle provide a mechanism for smoothing consumption and production in the face of shocks, but at the same time could expose a country to greater volatility as exogenous shifts in capital flows disrupt economic activity. Greater dependence on credit might make a country more vulnerable. In most of these cases the results on theoretical grounds alone are ambiguous; only a closer look at the data can reveal which effect dominates.

There are many candidates for explaining the higher volatility of developing economies. Among them, money growth, private capital flows, inflation, fiscal balances, and terms of trade are all more volatile in developing than in OECD countries. We consider these factors more systematically below.

Differences in Volatility across Regions

Not only are there significant differences in volatility across levels of economic development, there are also significant differences across regions. East Asia, especially the fast-growing group of six economies (the East Asia 6—Hong Kong, Indonesia, Malaysia, Singapore, Taiwan [China], and Thailand), has grown more quickly than the other developing regions (table 3). Until the recent crisis the East Asia 6 had achieved this growth without greater volatility: while the standard deviation of growth is higher in this group than in the developed world, it is considerably lower than the average for developing countries. In East Asia as a whole growth has been significantly higher, and volatility marginally higher, than the average for all developing countries.

This casual examination of the data corroborates much, but not all, of the conventional wisdom. East Asia, especially the East Asia 6, has far greater financial depth than do other developing regions. In fact, the East Asia 6 have slightly greater financial depth (as measured by credit to the private sector as a percentage of GDP) than the OECD countries, while the other developing regions have substantially less than the OECD countries. But the greater financial depth in East Asia comes at the cost of significantly higher volatility in financial activity, higher even than that in developed countries. The pattern is the same for money supply (M3) as a share of GDP. Both the level and the standard deviation of M3 as a share of GDP are higher among OECD than developing countries, and higher in East Asia (especially the East Asia 6) than other developing regions.

In trade openness East Asia, especially the East Asia 6, surpasses the developed countries. The reason is only partly size—that developed countries have

Table 3. Some Empirical Differences among Developing Regions

<i>Variable</i>	<i>East Asia</i>	<i>East Asia 6</i>	<i>Sub-Saharan Africa</i>	<i>South Asia</i>	<i>Latin America</i>
GDP growth	0.027	0.050	0.006	0.023	0.015
Standard deviation of GDP growth	0.062	0.031	0.064	0.033	0.050
Standard deviation of employment growth	0.119	0.125	0.131	0.119	0.088
Standard deviation of real wage index	2.108	2.390	1.916	2.074	1.592
Standard deviation of real wage growth	0.995	0.849	1.248	1.053	1.126
Standard deviation of fiscal balance	3.721	3.946	4.008	3.608	3.850
Credit to private sector/GDP	35.654	69.927	17.328	14.253	31.242
Standard deviation of credit to private sector/GDP	14.005	30.235	6.259	5.414	9.836
(Imports + exports)/GDP	94.071	135.318	68.509	44.759	80.979
M3/GDP	48.430	82.767	25.239	29.788	39.200
Standard deviation of M3/GDP	13.796	26.071	6.827	7.081	10.151
Standard deviation of inflation	0.107	0.116	0.154	0.085	0.232
Standard deviation of growth in M1	0.155	0.141	0.171	0.098	0.257
Private capital flows/GDP	2.260	2.105	1.619	0.504	2.067
Standard deviation of private capital flows/GDP	2.322	2.104	2.570	0.809	3.317
Standard deviation of terms of trade changes	0.068	0.064	0.145	0.114	0.117

Source: Authors' calculations.

smaller trade shares because they have larger economies—since this distinction holds when we add China to the East Asia 6. The ratio of trade to GDP is 118 for these seven economies, 123 for the East Asia 6, and 61 for OECD countries. Latin America is also more open than the OECD countries on average, while South Asia and Sub-Saharan Africa are less open. All these developing regions face greater volatility in terms of trade than do the developed countries. Among them, East Asia has by far the least volatility, and Sub-Saharan Africa the most.

There is little difference across the regions in the volatility of government activity (fiscal balance), but great differences in inflation and changes in money supply. Not surprising, the differences between OECD and non-OECD countries are driven by Latin America, where the volatility of both inflation and money growth are twice that in other developing regions.

Finally, the data cannot unambiguously justify the reputation of Latin America for wage rigidity. While Latin America has more stable wage levels than other regions, it has more volatility in wage growth than the OECD countries. Conversely, East Asia (especially the East Asia 6) has more stability in wage growth, although more volatility in wage levels.

Determinants of Volatility

Broadening our analysis to the variation in the entire cross-country sample, we find that openness to trade and volatility in terms of trade and in capital flows are all associated with volatility in per capita growth (table 4). All indicators of financial sector development are negatively associated with volatility, while volatility in money supply (M3) as a share of GDP is positively associated with volatility in growth. Wage and price flexibility and inflation variability are also associated with volatility in growth, as is variability in policy, whether fiscal or monetary.

To assess the relative impact of these factors, we regressed growth volatility (the standard deviation of the per capita growth rate) on a range of independent variables. Depending on the specification, our sample includes observations on 60–74 countries in a panel created by aggregating over the periods 1960–78 and 1979–97 (table 5; see the appendix for descriptive statistics of the sample).

Using a standard Hausman test, we found two variables to be endogenous—credit to the private sector and the standard deviation of private capital flows. These were

Table 4. Bivariate Correlations with Volatility in Growth of GDP Per Capita

Variable	Coefficient	t-statistic	P-value	R ²	Number of observations
Trade and financial openness					
Standard deviation of terms of trade changes	0.12006	3.284	0.001	0.073	139
(Imports + exports)/GDP	0.00013	2.040	0.043	0.023	177
Standard deviation of (imports + exports)/GDP	0.00106	3.661	0.000	0.072	176
Standard deviation of private capital flows/GDP	0.00237	3.834	0.000	0.086	159
Standard deviation of all capital flows/GDP	0.00214	3.280	0.001	0.062	166
Financial sector development					
Change in private credit/gross domestic investment	-0.17660	-8.633	0.000	0.315	164
Standard deviation of M3/GDP	0.00106	3.015	0.003	0.052	169
Stock market value traded/GDP	-0.04741	-1.819	0.072	0.036	92
Credit to private sector/GDP	-0.00041	-3.336	0.001	0.063	169
Long-term private debt issues/GDP	-0.17815	-2.166	0.037	0.113	39
Private bond market/GDP	-0.03451	-3.615	0.001	0.272	37
Public bond market/GDP	-0.02361	-2.626	0.013	0.165	37
Price variability and flexibility					
Inflation	0.03331	4.298	0.000	0.101	167
Standard deviation of real wage index	0.00368	1.654	0.101	0.025	109
Standard deviation of real wage changes	0.01127	3.481	0.001	0.106	104
Policy volatility					
Standard deviation of fiscal balance/GDP	0.00215	2.327	0.021	0.039	134
Standard deviation of inflation	0.04166	4.722	0.000	0.119	167
Standard deviation of money growth	0.06865	5.380	0.000	0.149	167
Other					
Per capita growth	-0.58696	-7.036	0.000	0.211	187
Dummy variable for OECD countries	-0.03515	-4.144	0.000	0.085	186

Source: Authors' calculations.

instrumented by a range of variables, including indicators for French or English legal origin, initial GDP per capita in each period, the urban share of the population, life expectancy, the standard deviation of terms of trade changes, indicators for oil and other commodity exports, and a measure of political stability (number of assassinations per million people). The set of instruments is both valid and sufficient: a Sargan test confirmed that the instruments are exogenous to the error in the second-stage regression, and an overidentification test ensured that the model is adequately identified. The likelihood ratio test for heteroskedasticity indicated that the errors differ systematically across countries. We correct for both heterogeneity and endogeneity (correlation between the regressors and the idiosyncratic error) using the method suggested by Baltagi (1995).⁸ The dummy variable for developing countries is significant only in the ordinary least squares specification, suggesting that we capture in the error correction models some of the structural factors that make developing economies more volatile.

Table 5. Determinants of Volatility in Growth

Variable	Ordinary least squares		Error correction two-stage least squares ^a					
Developing country dummy	0.008971	(2.37)**	0.001962	(0.41)	-0.001844	(-0.35)	0.000869	(0.17)
(Imports + exports)/GDP	0.000062	(2.43)**	0.000068	(1.98)*	0.000081	(2.16)**	0.000065	(1.84)**
Standard deviation of change in log real wage index	0.005861	(0.13)	-0.001801	(-0.04)				
Standard deviation of growth in M1	0.020729	(2.17)**	0.017451	(1.87)*	0.019222	(2.13)**		
Private capital flows/GDP	0.000133	(0.13)	0.000417	(0.31)	-0.000155	(-0.11)	0.000021	(0.02)
Standard deviation of private capital flows/GDP	-0.001136	(-0.88)	0.000230	(0.09)	0.000739	(0.36)	0.001303	(0.66)
Credit to private sector/GDP	-0.000200	(-1.25)	-0.000789	(-3.15)***	-0.000968	(-3.95)***	-0.000956	(-4.08)**
Credit to private sector/ GDP squared	0.000001	(0.98)	0.000004	(2.62)**	0.000004	(2.98)***	0.000004	(3.12)**
Intercept	0.028857	(4.20)***	0.048315	(5.41)***	0.056339	(6.08)***	0.056471	(6.35)**
F-test of all parameters (degrees of freedom)	(9,89)	(7.93)***	(9,59)	(8.51)***	(8,71)	(13.15)***	(7,73)	(10.52)**
Likelihood ratio test for heteroskedasticity (degrees of freedom) ^b	(59)	(709.62)***	(59)	(699.14)***	(71)	(930.42)***	(73)	(933.77)**
Sargan test (degrees of freedom) ^c			(9,88)	(0.60)	(9,114)	(0.96)	(9,118)	(1.16)
Overidentification test (degrees of freedom) ^d			(19)	(0.00)	(21)	(0.00)	(22)	(0.00)
Number of groups		60		60		72		74
Number of observations		98		98		124		128

* Significant at the 10 percent level.

** Significant at the 5 percent level.

*** Significant at the 1 percent level.

Note: T-statistics are in parentheses. We present robust standard errors for all results—that is, the variance of the error terms is allowed to vary systematically across countries but we assume that they are uncorrelated over time within countries. Because of the limited number of observations, it was necessary to drop certain variables, such as those relating to stock market development. We tried several other specifications related to our hypotheses but found additional variables to be insignificant (for example, inflation variability, external debt, terms of trade volatility, various measures of fiscal policy volatility, various measures of institutional development, the volatility of nominal exchange rates, and relevant interaction of nonlinear terms). We also checked for size effects, but they were insignificant. Standard errors corrected for clustering within countries.

a. Credit to the private sector, credit squared, and the standard deviation of private capital flows are endogenous. Instruments include dummy variables for French or English legal origin, initial GDP per capita in the period, the urban share of the population, life expectancy at birth, the standard deviation of terms of trade changes, dummy variables for oil exports and for other commodity exports, and the number of political assassinations per million population.

b. The null hypothesis is that the errors are homoskedastic across countries (Chi-squared).

c. The null hypothesis is that the instruments are not correlated with the residual (F-test).

d. The null hypothesis is that the instruments adequately identify the model (Chi-squared).

Source: Authors' calculations.

The results show that openness does expose a country to greater volatility in growth.⁹ Surprisingly, private capital flows and the standard deviation of private capital flows do not affect growth volatility in the multivariate instrumental variables regressions.¹⁰ Also surprisingly, volatility in real wages (indicating wage flexibility) does not seem to be significantly associated with volatility in output. On balance, we find evidence neither for the claim that wage and price rigidity is the cause of fluctuations nor for the claim that wage and price volatility increases output volatility through demand effects.

Our key result relates to the financial sector variables. Greater credit or a deeper financial system is significantly associated with less volatility in all specifications, but the relationship appears to be nonlinear. The squared term is significant and enters with a positive sign. While developed financial systems offer opportunities for stabilization, they may also imply higher leverage of firms and thus more risk and less stability. It appears that the consumption and production smoothing possibilities provided by a deep financial system might reduce growth volatility on average, particularly when shocks are small, *but up to a limit*. As the financial system grows relative to GDP, the increase in risk becomes more important and acts to reduce stability. There is a nonlinear relationship from the regression between credit to the private sector as a share of GDP and the standard deviation of growth, holding the other variables at their means. Very large financial sectors (which are of course rare) can magnify shocks to the economy, much like capital inflows and outflows can magnify boom-bust episodes.

Results from the same regression, but with initial GDP per capita rather than the dummy variable for developing countries, are basically similar (table 6). But openness loses significance. The use of an interaction term with initial GDP per capita reveals that openness may increase growth volatility but that this effect is significantly attenuated in richer countries.

Interestingly, variables such as the standard deviation of private capital flows and, in some specifications, private capital flows are not significant when variables representing the depth of the financial system are included. The reason may be that volatility in capital flows affects the economy primarily through its effect on the financial system and on financial variables. The significance of the credit variables is robust to various specifications.

Several regressions were run to test the significance of the financial variables when measures of institutional development or governance are included. International Country Risk Guide and Business Environment Risk Intelligence indicators of institutional development and indicators of democracy were all insignificant, and their inclusion did not affect the results (these results are available from the authors).

Downturns

As important as overall variability in output is, perhaps even more important are the large events—the periodic economic downturns that have long characterized market economies. To investigate what structural and institutional characteristics of economies

might explain downturns, we perform a probit analysis on a similar data set (table 7).¹¹ A downturn is defined as negative per capita growth, which takes on the value 1, while positive growth takes on the value 0. On average, countries experience declining real GDP roughly 20 percent of the time. Non-OECD countries are in a downturn 22 percent of the time, and OECD countries slightly more than 9 percent of the time.¹² Not surprisingly, countries that are growing faster have a lower probability of experiencing a downturn—the change in growth rate required for a recession is larger, and thus the shocks required to put the economy into recession are also larger.

The dummy variable for developing countries is significant—when we control for other variables, developing countries are far more likely to experience growth downturns than are industrial economies. This suggests again that there is something about the structure of developing economies that makes them more vulnerable to growth crashes, something that is not captured by the right-hand-side variables. Economies that are more open seem less likely to go into a growth downturn, even though they have greater variability in output as a result of a higher incidence of shocks. We are not sure what to make of this mixed result.

Greater financial sector depth, as measured by the ratio of credit to GDP, increases the likelihood of a downturn. But the squared term for financial sector depth is not significant in the probit regression. Equity markets have the predicted effect: such markets provide better risk diversification than do debt markets and thus reduce the

Table 6. Determinants of Volatility in Growth, with Initial GDP Per Capita

Variable	Model			
	(1)		(2)	
Initial GDP per capita	0.000311	(0.18)	0.005069	(2.07) **
(Imports + exports)/GDP	0.000055	(1.34)	0.000986	(2.69) ***
(Imports + Exports)/GDP times initial GDP per capita			-0.000107	(-2.56) **
Standard deviation of change in log real wage index	0.003414	(0.06)	-0.015017	(-0.30)
Standard deviation of growth in M1	0.017042	(1.69) *	0.017335	(1.70) *
Private capital flows/GDP	0.000294	(0.20)	-0.000759	(-0.51)
Standard deviation of private capital flows/GDP	0.001198	(0.35)	0.002708	(0.82)
Credit to private sector/GDP	-0.000846	(-2.84) ***	-0.000866	(-2.66) ***
Credit to private sector/GDP squared	0.000004	(2.50) **	0.000004	(2.28) **
Intercept	0.048115	(3.66) ***	0.006441	(0.36)
F-test of all parameters (degrees of freedom)	(9,59)	(52.48) ***	(10,59)	(53.20) ***
Likelihood ratio test for heteroskedasticity (degrees of freedom) ^a	(59)	(6774) ***	(59)	(7103) ***
Sargan test (degrees of freedom) ^b	(8,89)	(0.44)	(8,89)	(0.43)
Overidentification test (degrees of freedom) ^c	(19)	(0.00)	(20)	(0.00)
Number of groups		60		60
Number of observations		98		98

* Significant at the 10 percent level.

** Significant at the 5 percent level.

*** Significant at the 1 percent level.

Note: T-statistics are in parentheses. Standard errors corrected for clustering within countries.

a. The null hypothesis is that the errors are homoskedastic across countries (Chi-squared).

b. The null hypothesis is that the instruments are not correlated with the residual (F-test).

c. The null hypothesis is that the instruments adequately identify the model (Chi-squared).

Source: Authors' calculations.

Table 7. Probability of a Downturn

Variable	Model			
	a	b	c	d
Dummy variable for developing countries	0.519350 (3.48) ***	0.481812 (3.12) ***	0.536428 (3.67) ***	0.489651 (3.14) ***
Years since last downturn	0.019759 (1.56)	0.019814 (1.52)	-0.005692 (-0.48)	-0.003480 (-0.30)
Five-year moving average growth	-0.268670 (-6.30) ***	-0.263142 (-6.15) ***	-0.047695 (-1.55)	-0.052766 (-1.85) *
Credit to private sector/GDP	0.016989 (1.74) *	0.016645 (1.63)	0.018288 (1.79) *	0.019123 (1.79) *
Credit to private sector/GDP squared	-0.000071 (-0.89)	-0.000080 (-0.94)	-0.000072 (-0.86)	-0.000120 (-1.41)
Private capital flows/GDP	0.001520 (0.06)	-0.008840 (-0.39)	-0.033246 (-1.48)	-0.045774 (-1.89) *
Log change in real wages	-2.554385 (-2.94) ***	-2.629325 (-3.11) ***	0.530893 (0.87)	0.559962 (0.87)
Capital restrictions (Imports + exports)/GDP	-0.175419 (-0.94)	-0.163978 (-0.87)	-0.344664 (-1.80) *	-0.259596 (-1.45)
Stock market value traded/GDP	-2.194500 (-2.50) **	-2.241041 (-2.12) **	-3.855537 (-3.13) ***	-1.836889 (-2.33) **
Intercept	-0.982504 (-2.76) ***	-0.925632 (-2.61) ***	-1.070404 (-2.91) ***	-1.083129 (-2.69) ***
Chi-squared test of all parameters (10 degrees of freedom)	(124.99) ***	(101.79) ***	(53.03) ***	(47.51) ***
Chi-squared test of credit to private sector/GDP and credit to private sector squared	(6.71) **	(5.02) *	(9.43) ***	(3.19)
Log-likelihood	-198.59	-199.90	-229.58	-235.59
Number of countries	54	54	54	54
Number of observations	630	630	630	630

* Significant at the 10 percent level.

** Significant at the 5 percent level.

*** Significant at the 1 percent level.

Note: T-statistics are in parentheses.

a. Contemporaneous values of credit to the private sector, credit squared, private capital flows, and stock market value traded. Dependent variable = 1 if growth is negative from $t - 1$ to t .

b. Lagged values of credit to the private sector, credit squared, private capital flows, and stock market value traded. Dependent variable = 1 if growth is negative from $t - 1$ to t .

c. Contemporaneous values of credit to the private sector, credit squared, private capital flows, and stock market value traded. Dependent variable = 1 if growth is negative from t to $t + 1$.

d. Lagged values of credit to the private sector, credit squared, private capital flows, and stock market value traded. Dependent variable = 1 if growth is negative from t to $t + 1$.

Source: Authors' calculations.

economy's vulnerability to a downturn. The coefficient of the variable measuring the depth of the equity market has the predicted sign and is highly significant. This result, combined with the finding of a positive and marginally significant effect of credit on the likelihood of downturns, suggests that financial systems in which debt is more prominent than equity are more vulnerable to growth collapses. This again emphasizes the importance of financial variables in the analysis of volatility.¹³

Real wage flexibility (as measured by the change in the log of real wages) reduces the likelihood of a downturn, but this result is not robust to different lag specifica-

tions. That the length of an economic expansion does not have a statistically significant effect on the probability of a downturn may suggest that there is no mechanical business cycle. This confirms work on the United States, where Furman and Stiglitz (1998) have shown that there has been no regular business cycle (no dependence of the probability of a downturn on the length of the expansion) since World War II (see Stiglitz 1997).

Conclusion

This article can be thought of as a reexamination of the standard paradigm relating to economic stability. We began with the underlying hypothesis that a variety of dynamic effects and institutions important for understanding volatility have traditionally been omitted or underemphasized in standard economic models—and that some of the most important “omitted” variables are those relating to the financial sector. Analyses that neglect this broader range of variables may be badly off course in predicting an economy’s short-run performance.

Our empirical results support this hypothesis. Wage rigidity, at the center of traditional Keynesian analysis, seems on average to play little role in output variability. In contrast, financial variables are consistently significant in explaining both variability and the likelihood of a downturn. Of course, volatility will also differ across economies depending on their structure, the nature of the shocks they face, and the government’s policy regime. Openness and policy variability were also found to be significant determinants of growth volatility.

There are a host of microeconomic variables—such as firm net worth and cash flows—that we would have liked to bring into the empirical analysis of volatility. Unfortunately, data on such variables are scanty, available for only a few countries, over a limited period, and for a small sample of firms. Yet the theoretical analyses suggesting their importance are consistent with many aspects of the recent global financial crisis.

If correct, the results of our theoretical and empirical analyses have strong policy implications:

- Countries are often advised to make labor institutions more flexible—to allow more rapid lowering of real wages so that the demand for labor can adjust more rapidly to supply. But wage adjustments have aggregate demand effects, and the adverse effects of these may more than offset the positive effects arising from wage flexibility. Controlling for other variables, we find that on balance, real wage flexibility has neither negative nor positive effects on volatility.
- Countries are told that opening the capital account will allow risk diversification, stabilizing the economy. In fact, benefits on this score can be offset by the fact that capital movements are highly variable and can be highly procyclical, in some cases inducing downturns, in others exacerbating fluctuations arising from other sources. We find no evidence for either the stabilizing or the exacerbating role of capital flows. The sensible thing for

Appendix 1. Definitions and Sources of Data

<i>Variable</i>	<i>Definition</i>	<i>Source</i>
Growth	Annual change in log per capita income, constant 1995 U.S. dollars.	World Bank, World Development Indicators database
Real (industrial) wages	Total wage bill divided by total employment.	United Nations Industrial Development Organization database
Credit to private sector/GDP	Credit to the private sector includes loans, purchases of nonequity securities, and trade credits and other accounts receivable that establish a claim for repayment.	World Bank, World Development Indicators database
M1/GDP	M1 is the sum of currency outside banks and demand deposits other than those of the central government.	International Monetary Fund, International Financial Statistics database
M3/GDP	M3 is the sum of currency and deposits in the central bank (M0), plus transferable deposits and electronic currency (M1), plus time and savings deposits, foreign currency transferable deposits, certificates of deposit, and securities repurchase agreements (M2), plus travelers checks, foreign currency time deposits, commercial paper, and shares of mutual funds or market funds held by residents	International Monetary Fund, International Financial Statistics database
(Imports + exports)/GDP	Imports plus exports divided by GDP.	World Bank, World Development Indicators database
Inflation	Annual change in the consumer price index.	World Bank, World Development Indicators database
Private capital flows/GDP	Private capital flows consist of private debt and nondebt flows. Private debt flows include commercial bank lending, bonds, and other private credits; nondebt private flows are foreign direct investment and portfolio equity investment.	World Bank, World Development Indicators database
Terms of trade changes	Change in the export price index minus change in the import price index, weighted by openness.	World Bank, World Development Indicators database
Stock market value traded/GDP	Stock market value traded refers to the total value of shares traded during the period.	Beck, Demirgüç-Kunt, and Levine (1999)
French or English legal origin	Dummy variables.	World Bank data
GDP per capita	Constant 1995 U.S. dollars.	World Bank, World Development Indicators database
Urban share of population	Population in urban areas as share of total population.	World Bank, World Development Indicators database
Life expectancy at birth	Number of years a newborn would live if prevailing patterns of mortality at the time of its birth were to stay the same through its life.	World Bank, World Development Indicators database
Employment	—	World Bank, World Development Indicators database
Fiscal balance	Current revenue less total expenditure (central government).	World Bank, World Development Indicators database

Appendix 1. cont.

<i>Variable</i>	<i>Definition</i>	<i>Source</i>
Oil exporting country	Dummy=1 if the value of oil exports is more than half the value of total exports of goods and services.	World Bank data
Commodity exporting country	Dummy=1 if the value of primary commodity exports (other than oil) is more than half the value of total exports of goods and services.	World Bank data
Long-term private debt issues/GDP	Equity issues by the private sector divided by GDP.	Beck, Demirgüç-Kunt, and Levine (1999)
Private bond market/GDP	Total amount of outstanding debt securities issued by private entities divided by GDP.	Beck, Demirgüç-Kunt, and Levine (1999)
Public bond market/GDP	Total amount of outstanding debt securities issued by public entities divided by GDP.	Beck, Demirgüç-Kunt, and Levine (1999)
Assassinations per million	Political assassinations per million population.	Political Risk Services, <i>International Country Risk Guide</i>
Capital restrictions	Dummy=1 if restrictions exist on capital account payments.	International Monetary Fund, <i>Yearbook of Exchange Controls</i>

Appendix 2. Descriptive Statistics

<i>Variable</i>	<i>Mean</i>	<i>Standard deviation</i>
Volatility regressions (subset of 98 observations)		
Standard deviation of GDP per capita growth	0.034	0.016
(Imports + exports)/GDP	62.141	50.896
Standard deviation of growth in real wage index	0.043	0.033
Standard deviation of growth in M1	0.140	0.200
Private capital flows/GDP	1.096	1.575
Standard deviation of private capital flows/GDP	1.342	1.310
Credit to private sector/GDP	42.582	32.059
English legal origin	0.357	0.482
French legal origin	0.531	0.502
Initial GDP per capita	6,315.871	7,776.980
Urban population share	55.263	23.453
Life expectancy at birth	66.060	8.577
Standard deviation of terms of trade changes	0.084	0.054
Oil exporter	0.031	0.173
Other commodity exporter	0.143	0.352
Assassinations per million	0.038	0.161
Probits (all 630 observations)		
Downturn from t - 1 to t	0.138	0.345
Downturn from t to t + 1	0.137	0.344
Developing country	0.494	0.500
Length of expansion	8.121	7.377
Five-year moving average growth	3.216	2.611
Credit to private sector/GDP	57.812	30.401
Lagged credit to private sector/GDP	56.271	29.547
Private capital flows/GDP	1.123	2.935
Lagged private capital flows/GDP	1.001	2.717
Growth in real wage index	0.011	0.092
Capital restrictions	0.706	0.338
(Imports + exports)/GDP	61.602	30.565
Stock market value traded/GDP	0.097	0.199
Lagged stock market value traded/GDP	0.084	0.176

policymakers would be to devise new financial strategies that hedge against the risks of sudden outflows while maintaining access to finance.

- Openness enhances economic growth, and higher economic growth reduces volatility and thus the vulnerability to economic downturns (openness also has a direct effect in making downturns less likely).¹⁴ But we find that openness also contributes significantly to volatility in per capita GDP growth.

Standard macroeconomic models give short shrift to financial institutions, often seeming to suggest that the entire financial sector can be embedded in a money demand equation. Our analysis confirms that financial institutions play a central role in economic volatility and downturns—that financial depth (as measured by private credit as a share of GDP) reduces volatility up to a point, but that too much private credit can increase volatility. The financial sector can also exacerbate downturns, particularly if debt increases relative to equity.

Notes

1. The historical references are from Kindleberger (1978, pp. 212–13).

2. Traditional Keynesian theory focused on asymmetries in the adjustment of wages and prices; here we argue that asymmetries in the adjustment of real variables are every bit as important.

3. Technically, a firm is in bankruptcy only if its creditors have gone to court to seek redress or the firm has gone to court to seek protection from creditors. We use the term *distress* more generically to refer to situations in which the firm's net worth is negative or its cash flow (including what creditors are "voluntarily" willing to lend or roll over) is insufficient to meet its debt obligations.

4. These unpaid liabilities inhibit the activities of both firms and their creditors. The debt overhang is a liability to firms, yet it is not really an asset to financial institutions, which necessarily must take a conservative position in discounting the likelihood of being repaid.

5. The reason in theory is that imperfections in equity markets (which themselves can be explained by informational imperfections; few would question the hypothesis of limitations in equity markets for most developing countries) limit the extent to which risks can be shared and shifted, and agency problems in large corporations lead to incentive schemes that induce risk-averse behavior in managers (see, for example, Leland and Pyle 1977, Stiglitz 1982, and Greenwald and Stiglitz 1991). The evidence consists of a large catalogue of firm behaviors that are hard to reconcile with the standard neoclassical model with risk-neutral firms but are consistent with the theory of the risk-averse firm (see the discussion above as well as Stiglitz 1982 and Greenwald and Stiglitz 1991).

6. Studies also show that climatic, political, and terms of trade instability lowers growth rates (Guillaumont, Guillaumont, and Brun 1999 for Africa, Barro and Sala-i-Martin 1994, and Mendoza 1994). The Inter-American Development Bank (IADB 1995) estimated that the effects of greater volatility in terms of trade, the real exchange rate, and monetary and fiscal policy in Latin America have translated into a reduction in growth rates of around 1 percent a year.

7. For Latin America, the Inter-American Development Bank (IADB 1995) and Hausmann and Gavin (1996) have shown that external shocks (to the terms of trade and to capital flows) and the volatility of economic policy are associated with volatility in growth rates. The IADB report also argues that financial system weaknesses and exchange rate policies have been important determinants of growth volatility (pegging the exchange rate rather than choosing a more flexible regime tends to be associated with increased volatility in output).

8. This is a generalized two-stage least squares procedure using the within effects (standard deviations) and between effects (country means) as instruments on data transformed by a weighted average of the within and between variance components. The elements of the omega (weighting) matrix are computed as $\Omega^{-1/2} = (P/\sigma_{fe}) + (Q/\sigma_{be})$, where Q refers to the country variable means, P to the standard deviations, σ_{fe} to the standard errors from the fixed effects regressions, and σ_{be} to the standard errors from the between-effects regressions. See Over (1999) for further details.

9. This is consistent with the finding of Easterly and Kraay (1999) that small states are more volatile when they are more open. But openness still has a direct positive effect on mean growth that outweighs its effect through increasing volatility.

10. Countries with greater access to private capital at any time may be expected to have lower income volatility. When the estimation also includes values of private capital flows adjusted for errors and omissions, the results are unaltered.

11. Because data on stock market value traded are available only from 1970 on, this financial variable was excluded from the volatility equations, and the probit regressions were run on data from 1976 on (see the appendix for details).

12. This finding applies both to the complete sample of 170 countries and to the smaller subsamples used in the regressions.

13. External debt was tested as an explanatory variable but was not significant.

14. In the standard theoretical model greater openness induces greater efficiency, a one-time gain in productivity, but does not lead to sustained increases in economic growth. But the conventional wisdom, and much of the econometric literature, argue that openness has not only one-time efficiency effects but also long-term growth benefits, perhaps as a result of the discipline provided by enhanced competition, as a result of the increased awareness of new technologies, or as a result of the availability of a broader array of intermediate inputs. Endogenous growth models also predict a significant effect of openness on growth.

References

- Aizenman, J., and N. Marion. 1998. "Volatility and Investment: Interpreting Evidence from Developing Countries." *Economica* 66 (262): 157–79.
- Ball, L., N. G. Mankiw, and David Romer. 1988. "The New Keynesian Economics and the Output-Inflation Trade-Off." *Brookings Papers on Economic Activity* 1: 1–82.
- Baltagi, Badi H. 1995. *Econometric Analysis of Panel Data*. Chichester, N.Y.: Wiley.
- Barro, Robert J., and Xavier Sala-i-Martin. 1994. "Quality Improvements in Models of Growth." NBER Working Paper 4610. National Bureau of Economic Research, Cambridge, Mass.
- Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine. 1999. "A New Database on Financial Development and Structure." Policy Research Working Paper 2146. World Bank, Development Research Group, Washington, D.C.
- Bruno, Michael, and William Easterly. 1998. "Inflation Crises and Long-Run Growth." *Journal of Monetary Economics* 41 (1): 3–26.
- Corbo, V., S. Fischer, and S. Webb, eds. 1992. *Adjustment Lending Revisited: Policies to Restore Growth*. Washington, D.C.: World Bank.
- De Long, J. B., and Lawrence H. Summers. 1988. "How Does Macroeconomic Policy Affect Output?" *Brookings Papers on Economic Activity* 2: 433–80.
- Easterly, William. 1990. "Portfolio Effects in a CGE Model: Devaluation in a Dollarized Economy." In Lance Taylor, ed., *Structuralist Computable General Equilibrium Models for the Developing World*. Cambridge, Mass.: MIT Press.

- Easterly, William, and Aart Kraay. 1999. "Small States, Small Problems?" Policy Research Working Paper 2139. World Bank, Washington, D.C.
- Easterly, William, and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112 (November): 1203–50.
- Furman, J., and Joseph E. Stiglitz. 1998. "Economic Crises: Evidence and Insights from East Asia." *Brookings Papers on Economic Activity* 2: 1–135.
- Gavin, M., R. Hausmann, R. Perotti, and Ernesto Talvi. 1996. "Managing Fiscal Policy in Latin America and the Caribbean: Volatility, Procyclicality, and Limited Creditworthiness." Working Paper 326. Inter-American Development Bank, Office of the Chief Economist, Washington, D.C.
- Greenwald, B., and Joseph E. Stiglitz. 1991. "Examining Alternative Macroeconomic Theories." In Edmund Phelps, ed., *Recent Developments in Macroeconomics*. Brookfield, Vt.: Edward Elgar.
- Greenwald, B., Joseph E. Stiglitz, and A. Weiss. 1984. "Informational Imperfections in the Capital Market and Macroeconomic Fluctuations." *American Economic Review* 74 (2): 194–99.
- Guillaumont, P. J., S. Guillaumont, and Jean-François Brun. 1999. "How Instability Lowers African Growth." *Journal of African Economies* 8 (1): 87–107.
- Hausmann, R., and M. Gavin. 1996. "Securing Stability and Growth in a Shock Prone Region: The Policy Challenge for Latin America." Working Paper 315. Inter-American Development Bank, Office of the Chief Economist, Washington, D.C.
- Hellmann, Thomas, and Joseph E. Stiglitz. 2000. "Credit and Equity Rationing in Markets with Adverse Selection." *European Economic Review* (44) 2: pp. 281–304.
- Hellmann, Thomas, K. Murdock, and Joseph E. Stiglitz. 1996. "Deposit Mobilization through Financial Restraint." In N. Hermes and R. Lensink, eds., *Financial Development and Economic Growth*. New York: Routledge.
- IADB (Inter-American Development Bank). 1995. *Economic and Social Progress in Latin America: 1995 Report—Overcoming Volatility*. Baltimore, Md.: Johns Hopkins University Press.
- International Monetary Fund. Various years. *Yearbook of Exchange Controls*. Washington, D.C.
- Islam, N. 1995. "Growth Empirics: A Panel Data Approach." *Quarterly Journal of Economics* 110 (November): 1127–1170.
- Kane, E. 1990. "Incentive Conflict in the International Regulatory Agreement on Risk-Based Capital." NBER Working Paper 3308. National Bureau of Economic Research, Cambridge, Mass.
- Kindleberger, Charles P. 1978. *Manias, Panics, and Crashes: A History of Financial Crises*. New York: Basic Books.
- Leland, H. E., and D. H. Pyle. 1977. "Informational Asymmetries, Financial Structure, and Financial Intermediation." *Journal of Finance* 32 (2): 371–87.
- Mayer, Colin. 1988. "New Issues in Corporate Finance." *European Economic Review* 32 (June): 1167–83.
- Mendoza, E. G. 1994. "Terms-of-Trade Uncertainty and Economic Growth: Are Risk Indicators Significant in Growth Regressions?" International Finance Discussion Paper 491. Board of Governors of the U.S. Federal Reserve System, International Finance Division, Washington, D.C.
- Myers, S. C., and N. Maljuf. 1984. "Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have." *Journal of Financial Economics* 13 (June): 187–221.
- Newbery, D. M., and Joseph E. Stiglitz. 1982. "Risk Aversion, Supply Response, and the Optimality of Random Prices: A Diagrammatic Analysis." *Quarterly Journal of Economics* 97 (1): 1–26.

- Orszag, P., and Joseph E. Stiglitz. 1999. "Bankruptcy, Credit Constraints, and Economic Policy." World Bank, Washington, D.C.
- Over, M. 1999. "XTIVREG: Baltagi's Suggested Feasible Procedure for G2SLS and EC2SLS." World Bank, Development Economics Research Group, Washington, D.C.
- Political Risk Services. 1987. *International Country Risk Guide*. Syracuse, N.Y.
- Pritchett, Lant. 1998. "Patterns of Economic Growth: Hills, Plateaus, Mountains, and Plains." Policy Research Working Paper 1947. World Bank, Development Research Group, Washington, D.C.
- Ramey, G., and V. A. Ramey. 1995. "Cross-Country Evidence on the Link between Volatility and Growth." *American Economic Review* 85 (December): 1138–51.
- Stiglitz, Joseph E. 1982. "Information and Capital Markets." In W. F. Sharpe, C. Cootner, and P. H. Cootner, eds., *Financial Economics: Essays in Honor of Paul Cootner*. Englewood Cliffs, N.J.: Prentice Hall.
- . 1994. *Whither Socialism?* Cambridge, Mass.: MIT Press.
- . 1997. "The Long Boom? Business Cycles in the 1980s and 1990s." Paper presented at Georgetown University Macroeconomics Seminar, 4 September, Washington, D.C., and at Centre for Economic Policy Research conference The Long Boom, Stanford University, Stanford, Calif., 5 September.
- . 1999. "Toward a General Theory of Wage and Price Rigidities and Economic Fluctuations." *American Economic Review* 89 (2): 75–80.
- . Forthcoming. "Whither Reform?" In Boris Pleskovic and Joseph E. Stiglitz, eds., *Annual World Bank Conference on Development Economics 1999*. Washington, D.C.: World Bank.
- Stiglitz, Joseph E., and M. Uy. 1996. "Financial Markets, Public Policy, and the East Asian Miracle." *World Bank Research Observer* 11 (2): 249–76.
- Talvi, E. 1995. "Fiscal Policy and the Business Cycle Associated with Exchange Rate-Based Stabilizations: Evidence from Uruguay's 1978 and 1991 Programs." Working Paper 313. Inter-American Development Bank, Office of the Chief Economist, Washington, D.C.
- . 1996. "Exchange Rate-Based Stabilization with Endogenous Fiscal Response." Working Paper 324. Inter-American Development Bank, Office of the Chief Economist, Washington, D.C.
- United States, Executive Office of the President. 1997. *Economic Report of the President*. Washington, D.C.: Government Printing Office.
- World Bank. 1992. "Adjustment Lending and Mobilization of Private and Public Resources for Growth." Policy and Research Series, no. 22. Country Economics Department, Washington, D.C.
- . 1999. "Coping with the Crisis in Education and Health." *Thailand Social Monitor* (World Bank, Bangkok) 2 (July).

Creative Destruction and Development: Institutions, Crises, and Restructuring

Ricardo J. Caballero and Mohamad L. Hammour

Creative destruction, driven by experimentation and the adoption of new products and processes when investment is sunk, is a core mechanism of development. Obstacles to this process are likely to be obstacles to progress in standards of living. Underdeveloped and politicized institutions are a major impediment to a well-functioning process of creative destruction. They result in sluggish creation, technological “sclerosis,” and spurious reallocation of factors of production. These ills reflect the macroeconomic consequences of contracting failures in the presence of sunk investments. Recurrent crises are another major obstacle to creative destruction. The common inference that increased liquidations during crises result in increased restructuring is unwarranted. On the contrary, there are indications that crises freeze the restructuring process, an effect associated with the tight financial market conditions that follow a contraction. This productivity cost of recessions adds to the traditional costs of resource underutilization.

The world economy is undergoing momentous reorganization in the face of the development and large-scale adoption of information technologies. Alan Greenspan (1999) describes the United States’ recent experience of such reorganization:

The American economy, clearly more than most, is in the grip of what . . . Joseph Schumpeter . . . called “creative destruction,” the continuous process by which emerging technologies push out the old. . . . The remarkable . . . coming together of the technologies that make up what we label IT—information technologies—has begun to alter, fundamentally, the manner in which we do business and create economic value.

Ricardo J. Caballero is the Ford international professor of economics at the Massachusetts Institute of Technology and a research associate at the National Bureau of Economic Research. Mohamad L. Hammour is a research director at DELTA (a joint research unit of the National Center for Scientific Research, École Normale Supérieure, and École des Hautes Études en Sciences Sociales) and a research fellow of the Centre for Economic Policy Research. The authors are grateful to Abhijit Banerjee for his discussion of their paper and to two anonymous referees for their suggestions.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

This wave of restructuring is only the latest manifestation of creative destruction—by which the production structure weeds out unproductive segments; upgrades its technology, processes, and output mix; and adjusts to the evolving regulatory and global environment.

Ongoing restructuring is as relevant for the developing world as it is for economies at the leading edge of technology. In this article we draw on significant advances over the past decade in theoretical and empirical research on creative destruction to formulate propositions about the role and workings of this mechanism in the development process. Some of these ideas are firmly grounded in empirical evidence; others are no more than hypotheses, consistent with a combination of theoretical considerations and scattered evidence and deserving of systematic investigation in the future.

The rest of this article is organized into three sections. The first section reviews recent international evidence on gross job flows that supports the idea that creative destruction is a core mechanism of growth in market economics. The second section shows how underdeveloped and politicized institutions in the developing world are likely to constitute a major obstacle to a well functioning creative destruction process, and explores their consequences. The third section argues that recurrent crises in developing countries are a second major obstacle to creative destruction and reviews evidence indicating that such crises can freeze the restructuring process.

Creative Destruction and Economic Growth

The notion that creative destruction is at the core of economic growth in market economies goes back to Joseph Schumpeter (1942), who considered it “the essential fact about capitalism” (p. 83). Underlying any notion of restructuring is the assumption that choices of technology, output mix, and modes of organization are embodied in capital and skills. The resulting irreversibility of investment means that adjusting the production structure requires scrapping existing investments and replacing them with new ones. Conversely, if capital were perfectly malleable and skills fully generic, adjustment would be costless and instantaneous. Conceptually, it is the embodiment of technology combined with incessant opportunities to upgrade the production structure that places ongoing restructuring at the core of the growth process, regardless of whether the economy is a technological leader or a laggard.

Restructuring is closely related to factor reallocation. If investments need to be scrapped, it is because they are working with factors of production that must be freed up to combine with new forms of investment. In other words, restructuring generates a reallocation of factors in which technology is not embodied. This link has been exploited empirically to develop measures of reallocation that can be used as an index of restructuring. The most successful measures that have been developed are based on labor reallocation, although studies have attempted to look at other factors (see, for example, Ramey and Shapiro 1998).¹

The literature on gross job flows has constructed measures of aggregate gross job creation and destruction based on microeconomic data at the level of business units—plants or firms (see Davis and Haltiwanger forthcoming for an excellent sur-

vey). Gross job creation over a given period is defined as employment gains summed over all business units that expand or start up during the period; gross job destruction refers to employment losses summed over all units that contract or shut down. Although job flows are a useful indicator of restructuring, the link between the two is loose. Plant equipment and organization could be entirely upgraded with no change in the number of jobs; conversely, jobs could migrate from one location to another (such as for tax reasons) with no change in the activity performed.

From the many studies that construct measures of job flows for different countries, three features of the data have emerged that allow us to characterize the role of creative destruction in the growth process:

- Gross job creation and destruction are large, ongoing, and persistent.
- Most job flows take place within rather than between narrowly defined sectors of the economy.
- Reallocation of jobs from less productive to more productive business units plays a major role in industry-level productivity growth.

Gross Job Flows—Large, Ongoing, and Persistent

Job flows are generally large, both in high-income countries and in the few developing and transition economies for which we have data (table 1). An average of at least one in 10 jobs commonly turns over in a year. Job creation and destruction are simultaneous and ongoing. In the U.S. manufacturing sector in 1973–88, for example, the lowest rate of job destruction in any year was 6.1 percent, in the 1973 expansion, and the lowest rate of creation was 6.2 percent, in the 1975 recession (Davis, Haltiwanger, and Schuh 1996, table 2.1). Moreover, the bulk of those flows did not reflect temporary layoffs, which would not correspond to true restructuring. Data for several countries show high persistence rates for job creation and destruction (the percentage of newly created jobs that remain filled over the period, or of newly destroyed jobs that do not reappear) over one- and two-year periods (table 2). Overall, job flow data indicate extensive, ongoing restructuring.

Job Flows Largely within Rather Than between Sectors

To measure the creation and destruction that take place simultaneously beyond what is required to accommodate net employment changes, we define excess job reallocation as the sum of job creation and destruction minus the absolute value of net employment change. In data for several economies the share of excess job reallocation accounted for by employment shifts between narrowly defined sectors never exceeds 20 percent, and it is typically far smaller (table 3).

Two major factors seem to be behind reallocation within sectors: adjustment and experimentation. Several job characteristics that are important determinants of employment adjustment are not captured by output-based sectoral classifications. A job may be associated with outdated capital or skills (see, for example, Caballero and Hammour 1996b) or may have suffered a highly idiosyncratic disruption. Experimentation, in the

Table 1. Average Annual Gross Job Flows in Selected Countries

(percentage of employment)

Country	Period	Coverage	Employer unit	Job creation	Job destruction
<i>High income</i>					
Canada	1974–92	Manufacturing	Plant	10.9	11.1
Canada	1983–91	All employees	Firm	14.5	11.9
Denmark	1981–91	Manufacturing	Plant	12.0	11.5
Denmark	1983–89	Private sector	Plant	16.0	13.8
Finland	1986–91	All employees	Plant	10.4	12.0
France	1985–91	Manufacturing	Firm	10.2	11.0
France	1984–92	Private sector	Plant	13.9	13.2
Germany	1983–90	All employees	Plant	9.0	7.5
Italy	1984–93	Private sector	Firm	11.9	11.1
Netherlands	1979–93	Manufacturing	Firm	7.3	8.3
New Zealand	1987–92	Private sector	Plant	15.7	19.8
Norway	1976–86	Manufacturing	Plant	7.1	8.4
Sweden	1985–92	All employees	Plant	14.5	14.6
United Kingdom	1985–91	All employees	Firm	8.7	6.6
United States	1973–93	Manufacturing	Plant	8.8	10.2
United States ^a	1979–83	Manufacturing	Plant	10.2	11.5
United States ^a	1979–83	Private sector	Plant	11.4	9.9
<i>Middle and low income</i>					
Chile	1979–86	Manufacturing	Plant	13.0	13.9
Colombia	1977–91	Manufacturing	Plant	12.5	12.2
Estonia	1992–94	All employees	Firm	9.7	12.9
Morocco	1984–89	Manufacturing	Firm	18.6	12.1

a. Selected states. Based on data for employers covered by unemployment insurance.
Source: Davis and Haltiwanger forthcoming, table 3.2.

Table 2. Average Persistence Rates for Annual Job Flows in Selected Countries

(percent)

Country	Period	One-year horizon		Two-year horizon	
		Job creation	Job destruction	Job creation	Job destruction
Denmark	1980–91	71.0	71.0	58.0	58.0
France	1985–90	73.4	82.1	51.5	68.2
Netherlands	1979–93	77.9	92.5	58.8	87.3
Norway	1977–86	72.7	84.2	65.1	79.8
United States	1973–88	70.2	82.3	54.4	73.6

Note: The persistence rate refers to the percentage of newly created jobs that remain filled over the period, or of newly destroyed jobs that do not reappear.
Source: Davis and Haltiwanger forthcoming, table 3.6.

face of uncertain market prospects, technologies, cost structures, or managerial ability, appears to account for a large share of job flows (see, for example, Jovanovic 1982). This idea is supported by evidence from U.S. manufacturing and elsewhere that younger plants exhibit higher excess reallocation rates, even after a variety of plant characteristics are controlled for (see Davis and Haltiwanger forthcoming).

Traditional analyses of restructuring in the trade and development literature emphasize one dimension of creative destruction—major shifts between the main sectors of the economy. Much less attention goes to the multitude of creation and destruction

Table 3. Share of Excess Job Reallocation Accounted for by Employment Shifts between Sectors in Selected Countries

Country	Period	Classification scheme	Employer unit	Number of sectors	Average number of workers per sector (thousands)	Share of excess reallocation from shifts between sectors (percent)
<i>High income</i>						
Finland	1986–91	2-digit ISIC	Plant	27	49	6
France	1985–91	Detailed industry	Firm	600	37	17
Germany	1983–90	2-digit ISIC	Plant	24	1,171	3
Italy	1986–91	2-digit SIC private sector	Firm	28	322	2
Netherlands	1979–93	2-digit SIC	Firm	18	10	20
New Zealand	1987–92	2-digit ISIC	Plant	28	28	1
Norway	1976–86	5-digit ISIC manufacturing	Plant	142	2	6
Sweden	1985–91	2-digit ISIC	Plant	28	112	3
United States	1972–88	4-digit SIC manufacturing	Plant	448	39	13
<i>Middle and low income</i>						
Chile	1979–86	4-digit manufacturing	Plant	69	4	12
Colombia	1977–91	4-digit manufacturing	Plant	73	6	13
Morocco	1984–89	4-digit manufacturing	Plant	61	4	17

Note: ISIC is the International Standard Industrial Classification. SIC is the Standard Industrial Classification.
Source: Davis and Haltiwanger forthcoming, table 3.5.

decisions driven by highly decentralized, idiosyncratic factors and experimentation. Consideration of the role played by these underlying flows may shed new light on many conventional questions in development. For example, Levinsohn (1999) and Melitz (1999) argue that a significant benefit of trade reform arises through factor reallocation toward more productive firms. Similarly, Olley and Pakes (1996) find that deregulation in the U.S. telecommunications industry increased productivity predominantly through factor reallocation toward more productive plants rather than through productivity gains within plants. Another example, which we discuss later, is the effect of crises on restructuring activity and the costs of this effect in terms of productivity.

Job Reallocation a Major Factor in Productivity Growth

The function of large job flows within sectors and their relationship to productivity gains brings us to the third feature of the data, the evidence that factor reallocation is a core mechanism in productivity growth. In a careful study and survey of this issue Foster, Haltiwanger, and Krizan (forthcoming) examine U.S. manufacturing industries (classified at the four-digit level) over the 10 years from 1977 to 1987. They decompose industry-level multifactor productivity gains over the period into a within-plant term and a reallocation term. The within-plant term reflects productivity gains within continuing plants weighted by their initial output shares; the reallocation term reflects productivity gains associated with reallocation among continuing, entering, and exiting plants. They find that reallocation accounts on average for 52 percent of 10-year productivity gains. Entry and exit account for half this contribution: plants that exit during the period have lower productivity than continuing plants, while plants that enter catch up with continuing

plants only gradually, through learning and selection effects. Other studies of U.S. manufacturing based on somewhat different methodologies concur with the conclusion that reallocation accounts for a major part of the productivity growth within industries (Baily, Hulten, and Campbell 1992; Bartelsman and Dhrymes 1998).

It would be of great interest to know whether restructuring is as productive in developing countries as it is in the United States, but relevant studies are few and they raise methodological issues. Aw, Chen, and Roberts (1997) look at the question in Taiwan (China) and Liu and Tybout (1996) address it in Colombia, but both studies define the within-plant term of their productivity decomposition on the basis of a plant's average share over the period rather than its initial share.² As Foster, Haltiwanger, and Krizan (forthcoming) discuss, this approach tends to underestimate the contribution of reallocation between continuing plants.

Moreover, both studies conduct their decomposition over a horizon of less than 10 years: 5 years for Taiwan (China) and only 1 for Colombia. This reduces the contribution of entry, which takes place dynamically through learning and selection effects. This short-sample decomposition also is more sensitive to the cyclical nature of productivity, which can be expected to affect productivity growth mostly within plants. The studies show that reallocation accounts for 34 percent of average productivity gains in Taiwan (China) and almost zero in Colombia.³ Given the methodological differences between these studies and those in the United States, it is difficult to know whether these results imply that factor reallocation is less productive in those developing economies than it is in the United States.

The evidence of extensive, ongoing job flows that act as a major mechanism of productivity growth points to the centrality of creative destruction in the growth process. A corollary is that obstacles to creative destruction are likely to be obstacles to development and thus should be of central concern in development theory and policy. Such potential obstacles are the focus of the rest of this article.

Institutions and Restructuring

We have seen that the notion of restructuring presumes that investment is partly irreversible. When two factors of production enter into a production relationship, they develop a degree of specificity with respect to each other and to the choice of technology, in the sense that their value within this arrangement is greater than their value outside it. In the presence of specificity the institutional environment becomes critical. The reason is that irreversibility in the decision to enter a production relationship with another factor creates *ex post* quasi rents that need to be protected through *ex ante* contracting (Klein, Crawford, and Alchian 1978). If contracting ability is limited, the institutional environment determines the rules by which those quasi rents are divided. Poor institutions, by definition, prevent one of the parties to a transaction from getting the value of what it put in. This disrupts the broad range of financing, employment, and output sale transactions that underlie a healthy process of creative destruction.

We view institutional failure as the root obstacle to economic growth in the developing world (see Lin and Nugent 1995 for a broad review). This leads us to the pre-

sumption that poor institutions are likely to constitute a major disruption to creative destruction. To the extent that irreversibility of investment takes on an entirely new dimension in the presence of contracting difficulties, it becomes crucially important in the analysis of development.

In this section we propose a simple model of the distortions likely to affect the restructuring process and examine related empirical evidence. Our treatment of institutions is deliberately very generic. Our purpose is not to comment about specific arrangements, but to identify a common element that is likely to systematically affect creative destruction and that is shared by many examples of institutional failure—such as overly protective labor regulations, highly politicized and uncertain regulation of competition, and financial markets that lack transparency and investor protection.

Theoretical Considerations

We develop a basic model, based on Caballero and Hammour (1998b), that focuses on specificity in financing and employment relationships and its implications for aggregate restructuring. For this purpose we introduce three factors of production: capital, entrepreneurs, and labor. The specificity of capital with respect to entrepreneurs affects financing transactions; its specificity with respect to labor affects employment transactions. All three factors exist in infinitesimally small units. Entrepreneurs and labor have linear utility in the economy's unique consumption good, which we use as numeraire.

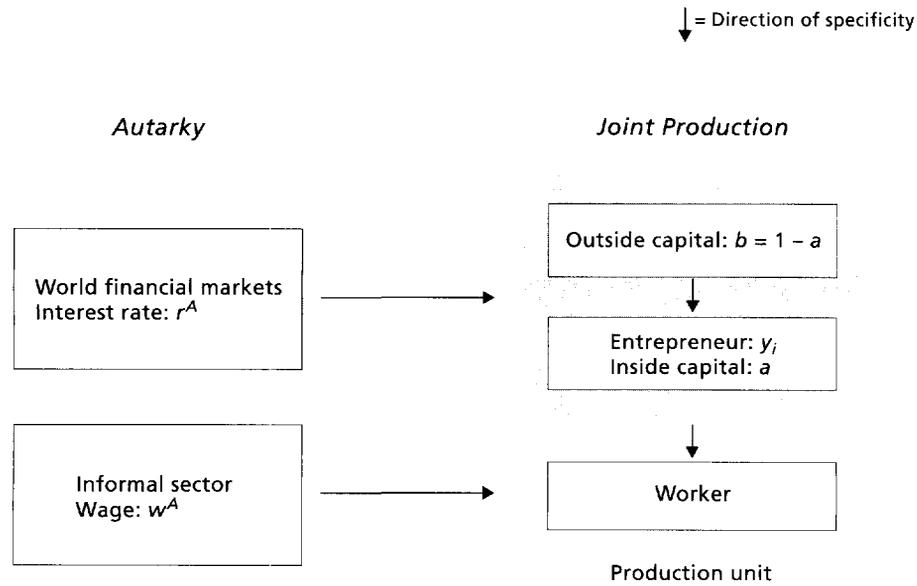
Contracting obstacles affect the possibility of economic cooperation. To capture their implications at a general level, we define for each factor two possible modes of production: autarky and joint production (figure 1). In joint production the three factors combine in fixed proportions to form production units. Each such unit is made up of a unit of capital, an entrepreneur i , and a worker. Each entrepreneur i has an innate level of skill that determines the production unit's productivity, measured by the amount, y_i , of the consumption good the unit can produce. Each entrepreneur also starts with a level of net worth $a_i \geq 0$ that can finance part of the unit's capital requirement. The remaining financing requirement, $b_i = 1 - a_i$, is provided by external financiers. We assume that workers start with zero net worth. Cooperation in joint production gives rise to investment specificity: once committed, capital is fully specific to the entrepreneur and the worker. It has no ex post use outside its relationship with them.

The autarky mode of production is free from investment specificity. If factors do not participate in a joint production unit, they can operate in the following autarky modes: Capital can be invested in the international financial markets at a fixed world interest rate $r^A > 0$ (A stands for autarky). An entrepreneur can also invest his net worth at the world interest rate. Workers can find employment in the informal sector at a wage w^A given by the informal sector labor demand function:

$$(1) \quad U = U(w^A), \quad U' < 0,$$

where U stands for informal sector employment.

Figure 1. Autarky and Joint Production



To analyze restructuring, we assume that the economy starts with preexisting production units as well as a supply of uncommitted factors of production. Events occur in three consecutive phases: destruction, creation, and production. In the destruction phase the factors in all preexisting units decide whether to continue to produce jointly or to separate and join the uncommitted factors. In the creation phase uncommitted factors either form new joint production units or remain in autarky. In the last phase production takes place and factor rewards are distributed and consumed. If the factors in a joint production unit separate after the creation phase, their only option is to move back to autarky.

Introducing preexisting units allows us to analyze destruction decisions. We assume that the units' productivity distribution is over the interval $y^o \in (0, y^{max})$ and, for simplicity, that it has negligible mass. The supply of uncommitted factors is as follows: The supply of capital is unlimited. The supply of entrepreneurs with any given productivity $y \in (0, y^{max})$ is also unlimited, but not all entrepreneurs have positive net worth. We assume that entrepreneurs with positive net worth are distributed according to a uniform density $\phi > 0$ for each productivity level and that they all have sufficient funds to fully finance a production unit ($a_i > 1$). The aggregate mass of labor is one, so that employment in joint production is given by

$$(2) \quad L = 1 - U(w^A) .$$

EFFICIENT EQUILIBRIUM. We first derive the economy's efficient equilibrium conditions, which would arise if agents had perfect contracting ability. We restrict ourselves

to parameter configurations that result in an interior equilibrium ($0 < L < 1$). On the creation side, since the supply of entrepreneurs with the highest productivity y^{max} is unlimited and the autarky return on capital is r^A , labor's autarky wage must satisfy

$$(3) \quad w^{A*} = y^{max} - r^A.$$

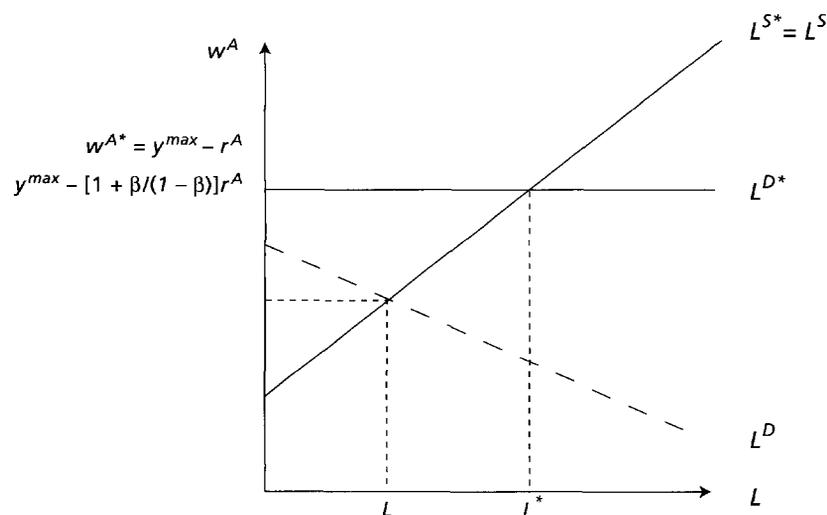
(The asterisk denotes efficient equilibrium values.) Any wage below this value would induce infinite joint production labor demand; any wage above this value would induce zero demand. The labor demand and supply system given by equations 2 and 3 determines the efficient equilibrium creation of joint production units, as illustrated in figure 2. The joint production rewards for capital and labor are equal to their autarky rewards, and the reward for entrepreneurs is zero because of their unlimited supply.

On the destruction side, scrapping the capital invested in a preexisting unit frees up a unit of labor. Efficient exit will therefore affect all units with productivity levels

$$(4) \quad y^o < w^{A*}.$$

INCOMPLETE-CONTRACTS EQUILIBRIUM. Because of investment specificity, implementing the efficient equilibrium requires a contract that guarantees capital in joint production its ex ante opportunity cost r^A . The contracting incompleteness we introduce is due to the inalienability of human capital, which renders unenforceable any contract clause that removes the right of the entrepreneur or worker to walk away from the joint production relationship ex post (see Hart and Moore 1994). This affects both the employment transaction between labor

Figure 2. Efficient and Incomplete-Contracts Equilibria



and capital and the financing transaction between the entrepreneur and external financiers.

In the employment relationship we assume that the worker deals with the entrepreneur and his financier as a single entity.⁴ If production unit i has productivity y_i , its associated specific quasi rent, s_i , is the difference between the unit's output and its factors' ex post opportunity costs:

$$(5) \quad s_i = y_i - w^A,$$

given that the worker moves to autarky if he leaves the production unit. Consistent with the Nash bargaining solution for sharing the unit's output, we assume that each party gets its ex post opportunity cost plus a share of the surplus s_i . If $\beta \in (0,1)$ denotes labor's share,

$$(6) \quad w_i = w^A + \beta s_i \quad \text{and} \quad \pi_i = (1 - \beta)s_i,$$

where w_i denotes the rewards of labor and π_i the rewards of capital. The contracting problem adds a rent component βs_i to wages.

In the financing relationship the associated specific quasi rents correspond to the full profit, π_i , because the ex post outside options of the entrepreneur and external financiers are worthless. Again, because of the inalienability of human capital, no contract can prevent the entrepreneur from threatening to leave the relationship. Any contract can be renegotiated according to the Nash bargaining solution, which gives the entrepreneur a share $\alpha \in (0,1)$ of π_i and the external financier a share $1 - \alpha$. The production unit's outside liability can therefore never exceed

$$(7) \quad r^A b_i \leq (1 - \alpha)\pi_i.$$

This financial constraint places a lower bound on the net worth, $a_i = 1 - b_i$, that the entrepreneur needs to start a project, which can be written as

$$(8) \quad a_i \geq 1 - (1 - \alpha)(1 - \beta)(y_i - w^A)/r^A$$

based on equations 5 and 6. We assume that a is large enough so that equation 8 requires positive net worth when $y_i = y^{max}$. This implies that only entrepreneurs with positive net worth can enter joint production, in which case we have assumed that they have enough funds to fully finance a production unit.

We now solve for the incomplete-contracts equilibrium conditions. In the creation phase an entrepreneur able to finance a production unit will find it profitable to do so if

$$(9) \quad \pi_i \geq r^A,$$

which, given equations 5 and 6, is equivalent to

$$(10) \quad y_i \geq w^A + [1 + \beta/(1 - \beta)]r^A.$$

Because of the rent component in wages, capital behaves as if it faced a world interest rate higher than r^A . The joint production demand for labor is given by the mass of entrepreneurs whose productivity satisfies equation 10 and who can finance a production unit:

$$(11) \quad L = \phi [y^{max} - w^A - [1 + \beta/(1 - \beta)]r^A].$$

Together with equation 2 for the supply of labor, this equation determines the incomplete-contracts equilibrium level of L . As illustrated in figure 2, labor demand (equation 11) under incomplete contracts falls below its efficient economy counterpart (equation 3). This occurs both because of labor market rents (which shift the curve down vertically) and because of the financial constraint (which rotates the curve down around its vertical axis intercept). In the incomplete-contracts equilibrium joint production employment and autarky wages are lower than their efficient equilibrium counterparts:

$$(12) \quad L < L^* \quad \text{and} \quad w^A < w^{A*}.$$

In the destruction phase a worker who leaves a preexisting production unit will find employment in joint production with probability L , in which case we denote his expected wage by $E(w)$, and will remain in autarky with probability $(1 - L)$, in which case his wage will be w^A . The exit condition is therefore⁵

$$(13) \quad y^o < L E(w) + (1 - L) w^A.$$

CHARACTERIZATION OF EQUILIBRIUM. We now characterize the general equilibrium consequences of incomplete contracting. The imbalances we describe constitute a highly inefficient macroeconomic solution to the unresolved microeconomic contracting problems. The first are highly suggestive of the experience of developing countries but pertain only indirectly to restructuring.

- *Reduced cooperation.* At the purely microeconomic level it is well known that limited contracting ability hampers cooperation. We have seen that under limited contracting ability even joint production projects with positive value may not be undertaken because workers (equation 6) or entrepreneurs (equation 7) can capture rents beyond their ex ante opportunity costs.
- *Underemployment.* As we have seen in the discussion of equation 12, joint production is characterized by underemployment ($L < L^*$), which is an equilibrium consequence of obstacles to cooperation in the financial and labor markets. In partial equilibrium, rent appropriation reduces the joint production return on capital. To restore this return to the level r^A required by world markets, fewer joint production units are created, informal sector employment bal-

loons, and the opportunity cost component w^A of wages falls (equation 6). The extent of underemployment generally depends on the supply elasticity of the factor that suffers from specificity, which we assume here to be infinite.

The counterpart of underemployment in joint production is an overcrowded informal sector ($U > U^*$). The informal sector becomes overcrowded because we have assumed no need for contracting in autarky. We view the informal sector as one in which transactional problems are less severe because there is less need for cooperation with capital (because of low capital intensity or constant returns, because employment regulations can be evaded, and so on).⁶

- *Market segmentation.* In the incomplete-contracts equilibrium both the labor market and the financial market are segmented. There are workers and entrepreneurs in autarky who would prefer to move into joint production but are constrained from doing so. Put another way, those two factors earn rents in joint production. It is easy to see that the rent component of joint production wages in equation 6 is positive, raising the wages above the informal sector wage.⁷ But the presence of rents does not entail high wages—quite the contrary. Joint production wages are lower under incomplete contracts than in the efficient economy. To see this for any production unit i , substitute $\pi_i = y_i - w_i$ into equation 9 to get

$$(14) \quad w_i \leq y_i - r^A \leq y^{max} - r^A = w^{A*}.$$

The rent component of wages arises as a result of depressed wages in the informal sector, not because of high wages in joint production. Similarly, from equation 10 it is clear that an entrepreneur with intramarginal productivity y_i earns positive rents equal to $y_i - w^A - [1 + \beta/(1 - \beta)] r^A$ associated with the scarcity of internal funds. Those rents would not arise in an efficient equilibrium.

We now turn to the characteristics of equilibrium that pertain directly to restructuring. The first three properties characterize the amount of equilibrium creation and destruction of joint production units; the last two characterize the quality of restructuring, understood as the net gain that results from it.

- *Depressed creation.* Since creation in this economy is equal to $L < L^*$, it follows that the equilibrium rate of creation is depressed compared with creation in the efficient economy.
- *Sclerosis.* The joint production structure suffers from sclerosis, in the sense that some production units survive that would be scrapped in an efficient economy. To see this, compare the efficient and incomplete-contracts exit conditions (equations 4 and 13). Since $w^A < w^{A*}$ was shown in equation 12 for autarky and $w_i < w^{A*}$ in equation 14 for joint production, it is clear that cost pressures to scrap are lower in the incomplete-contracts than in the efficient equilibrium. Sclerosis is thus a result of the underutilization and low productivity of labor. Sluggish creation and sclerosis can be a heavy drag on aggregate productivity.
- *Unbalanced restructuring.* Destruction is excessively high compared with the depressed rate of creation. To see this, note that the private opportunity cost

used in equation 13 for exit decisions is higher than the social shadow wage w^A of labor. The reason is that it is possible to capture a rent component in wages, which distorts the private opportunity cost of labor upward. That the economy exhibits both sclerosis and excessive destruction may appear paradoxical. In fact, the sclerosis reflects a comparison with the efficient equilibrium, the excessive destruction reflects a comparison between private and social values within the incomplete-contracts equilibrium. The imbalance of gross flows is closely related to the presence of rents and market segmentation. In Caballero and Hammour (1996a) we argue that this imbalance sheds light on the nature of employment crises in developing countries.

- *Scrambling.* In the efficient economy only the most productive entrepreneurs, with $y = y^{max}$, are involved in joint production. If the number of such entrepreneurs had been insufficient, others would have been brought in according to a strict productivity ranking. On the creation side an efficient process should result in implementation of the most productive projects. This ranking is scrambled in the incomplete-contracts equilibrium because another characteristic of the entrepreneur—net worth—comes into play. This tends to reduce the quality of the churn, as the same volume of scrapping and reinvestment will result in a smaller productivity gain.
- *Privately inefficient separations.* Another important consequence of contracting difficulties, although one that we have not incorporated in our model, is the possibility of privately inefficient separations. Such separations can arise as a result of factors similar to those that make creation privately inefficient—by constraining agents from starting positively valued projects. For example, assume that a production unit goes through a period of negative cash flow that must be financed if the unit is to remain in operation. Continuation investment would help preserve the unit's specific capital and is therefore itself specific and subject to a financial constraint. When the financial constraint is binding, destruction can be privately inefficient and result in losses for the owners of both labor and capital (for details see Caballero and Hammour 1999). This gives rise to another factor that reduces the quality of restructuring, since it generates spurious churn with little payoff in productivity gains. Moreover, once we admit the possibility of private inefficiency on the destruction margin, factors other than productivity may affect destruction decisions and also scramble the productivity ranking on the exit margin.

POLITICAL ECONOMY. Although in our model contracting incompleteness is based on the inalienability of human capital, it can be due to a variety of other factors. In particular, the legal and regulatory framework can be a source of factor specificity and provide the institutional framework that determines the division of specific rents. Legal restrictions on employee dismissals, for example, would effectively make capital partly specific to labor in the joint production relationship. Moving beyond an exogenous view of institutions, we look at some of the underlying causes of institutional obstacles to efficient restructuring.

Institutions serve two distinct functions: efficiency and redistribution. It is naive to think that markets can generally function properly without an adequate institutional framework. We have seen that the basic principle that determines institutions in their efficiency role is that each factor ought to get out the social value of what it put in—absent any externalities, its *ex ante* terms of trade. But it is equally naive to think that such institutions, being partly determined in the political arena, will not also be used as an instrument in the politics of redistribution. A poor institutional framework is the result of underdeveloped contracting and regulations combined with overly powerful political interest groups that have tilted the institutional balance excessively in their favor.

By displacing technologies and skills, creative destruction threatens a variety of incumbent interests and therefore can itself give rise to political opposition and endogenous institutional barriers. Mere uncertainty about the impact of restructuring can prop up opposition (Fernandez and Rodrik 1991). Mokyr (1992) discusses many historical examples of resistance to technology adoption, perhaps the most well known of which is the 19th-century Luddite movement in Britain (Thomis 1972). The resistance can range from mere neglect of urgent institutional reform to active barriers affecting trade, competition, regulation, the size of the government sector, and the aspects of financial and labor markets that we focus on in our model.

But protection of labor or other factors characterized by relatively inelastic supply can backfire, resulting in large-scale underemployment and internal segmentation between those who end up benefiting from protection and those who do not. This pitfall is worth highlighting as several Latin American economies (such as Argentina and Chile) revise their labor codes in the context of ever-increasing globalization and expanding options for external capital (see Caballero and Hammour 1998a).

A Look at Available Evidence

We have made a theoretical argument that poor institutions generally result in a stagnant and unproductive process of creative destruction. If institutional failure is considered the fundamental illness of the developing world, sclerosis and a low-quality churn could be presumed to be prevalent. Although this presumption is consistent with the low productivity in developing countries, it would be preferable to find more direct evidence for it from job flows. At first sight the data in table 1 do not seem to support sclerosis. Job flows in the few developing countries for which we have data are similar in size to those in high-income countries—if not larger (see Tybout 2000). But there are several powerful reasons why this evidence cannot be taken at face value.

MEASUREMENT ISSUES. Lack of uniformity in job flow measures may undermine their comparability across countries. Table 1 highlights major differences in sample coverage (manufacturing, the private sector, or all employees) and the basic employer unit (the plant or the firm). Other important differences are more difficult to trace, most notably the difficulties in linking observations longitudinally in the face of

ownership or other changes. For example, Contini and Pacelli (1995) report that attempts to correct Italian data for spurious births and deaths reduce job flows by about a fifth (see Davis and Haltiwanger forthcoming, table 3.2).

INDUSTRIAL STRUCTURE AND EMPLOYER CHARACTERISTICS. The magnitude of job flows varies systematically with industrial structure and employer characteristics. Davis and Haltiwanger (forthcoming) show that the industry pattern of job reallocation intensity is quite similar across countries. A regression of reallocation on industry fixed effects (with industries classified at the two-digit level) for pooled Canadian, Dutch, and U.S. data yields an R-squared of 48 percent. Although we are not aware of any systematic investigation of this issue, we would expect to find that developing country employment is heavily biased toward light industries with relatively low investment specificity and, typically, a fast turnover rate. This type of restructuring, with small reinvestment requirements, can be expected to yield commensurately small productivity gains. Moreover, it may even be an indication that developing countries avoid industries in which restructuring is expensive, rather than a sign of their ability to restructure.

Davis and Haltiwanger (forthcoming, figure 4.1) also summarize evidence from seven countries that shows that job reallocation rates fall significantly with employer size. The bias in the size distribution in developing countries toward small plants is dramatic compared with that in high-income countries (see, for example, Tybout 2000, table 1). This bias by itself predicts much larger job flows in developing economies. The contribution to productivity by this type of reallocation requires close interpretation. If small plant size is closely related to the light-industry bias with little technological specificity, the benefits of restructuring may be small. Moreover, if small plant size is associated with greater financial fragility, some of the turnover may be privately inefficient and unproductive.

RESTRUCTURING REQUIREMENTS. Given the catching up that developing economies have ahead of them, they could be expected to have significantly higher investment and restructuring requirements than industrial economies. The extraordinary turnover rates of Taiwanese firms may be a case in point. Aw, Chen, and Roberts (1997), in their study of Taiwanese manufacturing industries, report that new entrants over the previous five years accounted for a third to a half of industry output in 1991—compared with 14–19 percent in the United States, 15–16 percent in Chile, and 18–21 percent in Colombia. The high turnover rates in Taiwan (China) suggest that absent major impediments, developing countries have the potential to attain much higher restructuring rates.

Another useful natural experiment can be found in the transition to market economies in Eastern Europe (see Haltiwanger and Vodopivec 1997). In Estonia in 1992–94 annual job creation rates averaged 9.7 percent, and annual job destruction rates 12.9 percent—within the range observed in OECD economies (see table 1). What is striking is that these rates coincided with a period of momentous reforms—Estonia, one of the most radical reformers in the region, implemented major reforms in 1992. Between 1989 and 1995 employment by private enterprises rose from 2 percent of the total to 35 percent, and the share of establishments with more than 100 employees fell

from 75 percent to 46 percent. In this context the observed job flows in Estonia were disappointingly low—not surprising, given the major institutional deficiencies faced by transition economies.

PRODUCTIVITY. So far our discussion has been limited to the volume of the churn. Our theoretical discussion pointed to factors—privately inefficient separations and scrambling in the productivity ranking of entering and exiting units—that reduce the quality of those flows. In principle, sclerosis is consistent with large flows if those flows are relatively unproductive. The quality of the churn can be measured by an accounting exercise (like that discussed in the previous section) that accounts for the aggregate productivity improvements associated with job flows. As noted in our discussion of the results of studies in Colombia and Taiwan (China), methodological issues do not allow direct comparison with results for the United States. Just as important, those studies do not account for the scrapping and reinvestment costs of restructuring. When a firm exits and is replaced by an entrant with higher productivity, the cost of scrapping investments in the exiting firm and reinvesting in the entrant needs to be accounted for. This is particularly important in comparisons of high- and low-income economies, when employment in the low-income economies is biased toward light industries and other modes of production with low reinvestment costs.

* * *

It seems safe to conclude that cross-country comparisons based on raw job flow data are unlikely to provide conclusive evidence on the efficiency of restructuring. A more structural empirical approach is needed that addresses the types of issues discussed above. From this point of view the empirical literature is still in its infancy.

Crises, Recovery, and Productivity

Recurrent crises in developing economies have large welfare consequences. Some of these consequences are immediately apparent. Others are manifested over time and are thus often underappreciated. A potentially important example of the second type is the disruptive effect that crises can have on the restructuring process. In this section we report evidence that leads us to conjecture that crises slow restructuring. If this is true, and given our presumption of sclerosis in the production structure, crises are even costlier than their immediate impact on unemployment and other aggregate indicators might suggest.

The most noted impact of contractions on restructuring is a sharp increase in liquidations. Consider Chile's debt crisis in the early 1980s (figure 3). The job destruction rate in manufacturing exceeded 22 percent in 1981. Sharp increases in liquidations during recessions have also been documented for other countries.⁸ But it would be wrong to infer that the concentration of liquidations during crises means that crises accelerate restructuring. This view was influential among pre-Keynesian "liquidationists"—such as Friedrich Hayek, Arthur Pigou, Lionel Robbins, and

Joseph Schumpeter—who saw liquidations in a positive light as the main function of recessions (see De Long 1990).

Although few economists today would take such an extreme position, many see increased factor reallocation as the “silver lining” of recessions. Liquidations are viewed as a prelude to much-needed restructuring. Under the presumption of technological sclerosis due to poor institutions, increased restructuring can be beneficial. A variety of liquidationist arguments were advanced during the Asian crisis, for example, in connection with the reorganization of Korean *chaebol*.

Although there seems to be some truth to the notion that recessions facilitate reorganization in politics and institutions, the relationship between liquidations and restructuring is much less obvious in the production structure. Jobs lost during recessions typically feed into unemployment or into underemployment in the informal sector, not directly into increased creation—the phenomenon we referred to as unbalanced restructuring. The question is whether increased liquidations ultimately lead to increased restructuring.

To address this question, one needs to examine the cumulative impact of a recessionary shock on creation and destruction—not only the effect of the crisis at impact, but also how the recovery materializes. Figure 4 shows three scenarios that are consistent with a given unemployment recession that starts with a spike in liquidations (top panel). The three scenarios correspond to cases in which the recession results cumulatively in increased, unchanged, or decreased restructuring (bottom three panels).

Figure 3. Gross Manufacturing Job Flows in Chile, 1979–85



Source: Roberts 1996, table 2.2.

We examine this question empirically in Caballero and Hammour (1999) using data from the U.S. manufacturing sector. In gross job creation and destruction time series constructed by Davis and Haltiwanger (1992) for U.S. manufacturing, there is a sharp rise in destruction at the onset of each recession, but a much more muted fall in creation (figure 5). Although this asymmetry between creation and destruction may be less strong in other sectors or when the economy is subject to different types of shocks, this evidence confirms the long-held view that liquidations are highly concentrated in recessions.

But does the evidence support increased restructuring following recessions? To examine the cumulative impact of a recessionary shock on creation and destruction, we ran a simple one-factor regression and calculated impulse-response functions (figure 6). Surprisingly, recessions seem to reduce the amount of restructuring in the economy. This finding of a “chill” following recessions is significant and robust in several dimensions, including to the introduction of a second, reallocation shock. Given the limitations of the data, our conclusion can be only tentative. But the evidence does not support the prevailing views that recessions cause increased restructuring.

Why would recessions freeze the restructuring process? Based on the model we develop in Caballero and Hammour (1999), our interpretation is that the main under-

Figure 4. Cumulative Effect of Crises on Restructuring: Three Scenarios

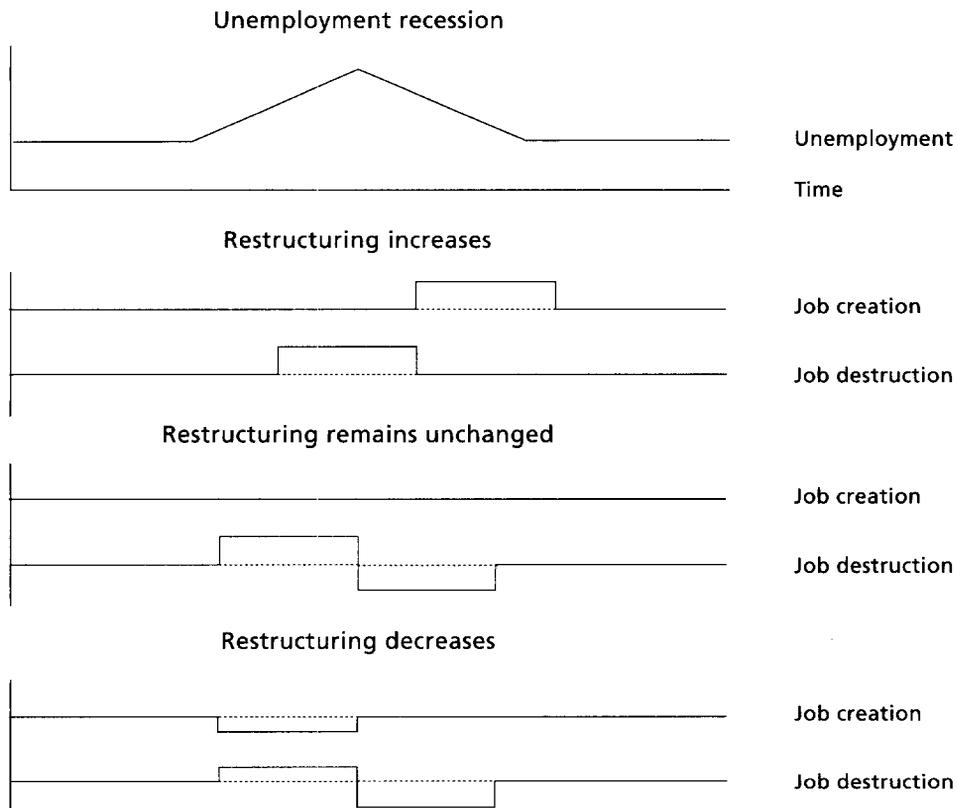


Figure 5. Gross Manufacturing Job Flows in the United States, 1972–92

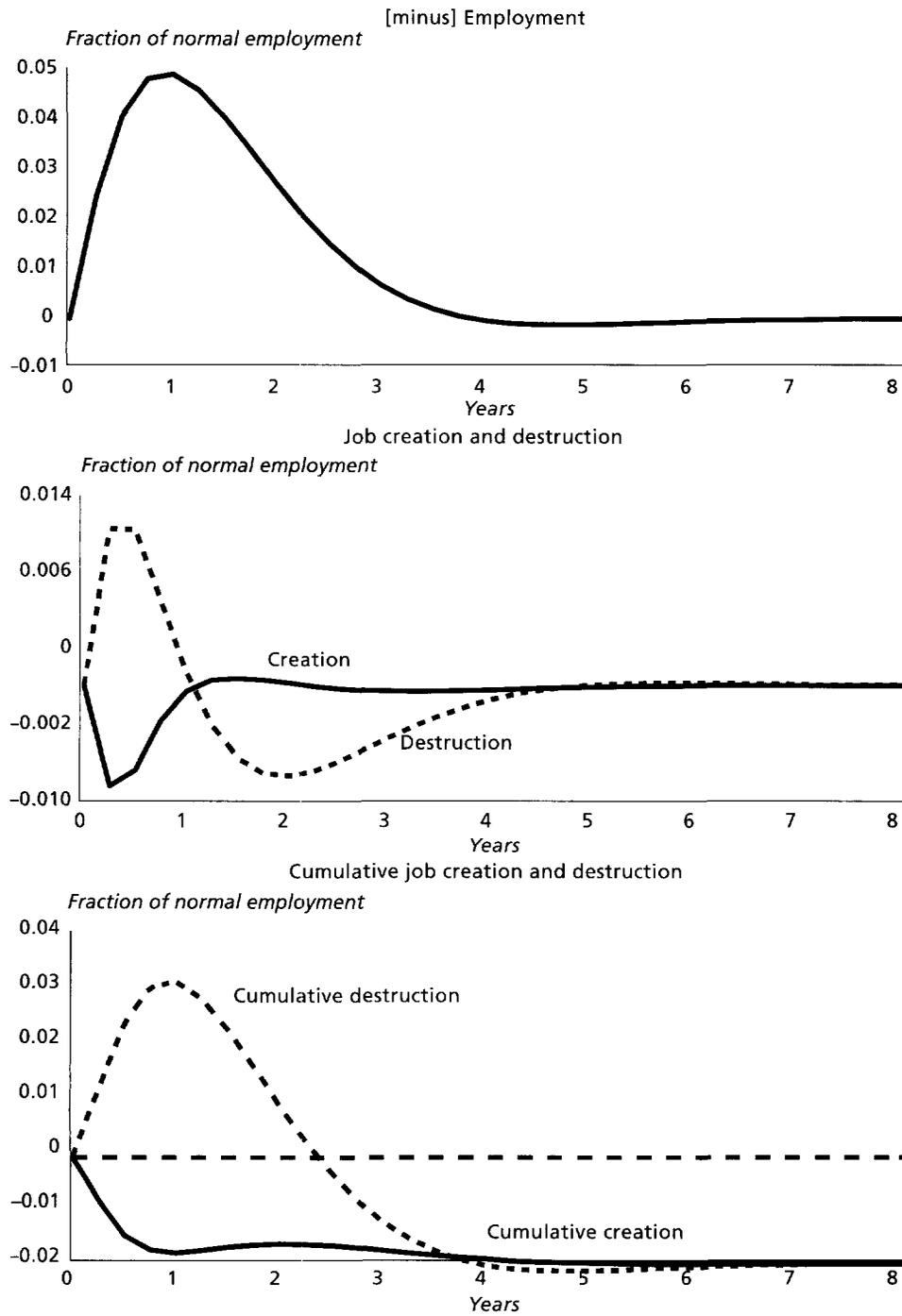
Source: <http://www.bsos.umd.edu/econ/haltiwanger/download.htm>.

lying factors are financial constraints—again a case of institutional failure. Liquidations and bankruptcies make the news, but recessions also squeeze the liquidity and financial resources needed to create new, more advanced production units. The competitive pressure from new production units therefore lessens, allowing low-productivity incumbents to survive more easily. The scarcity of financial resources during the recovery limits the socially useful transfer of resources from low- to high-productivity units.⁹

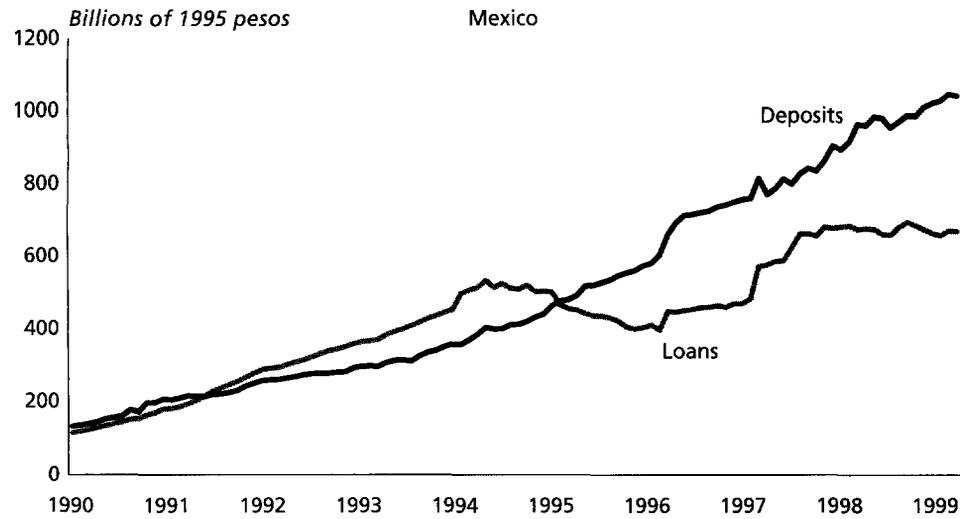
While lack of data preclude us from reproducing this analysis for a developing economy, it is plausible that the same phenomenon characterizes crises in developing economies.¹⁰ The liquidity contractions in those economies are more marked, however, and their depressing effect on creation during the recovery is likely to be even stronger. In Argentina and Mexico, for example, a severe credit crunch followed the “tequila” crisis of the mid-1990s (figure 7). Loans to the private sector not only recovered very gradually after the crisis, but did so more slowly than deposits.¹¹

Thus even though direct evidence is lacking, it is likely that crises constitute another major obstacle to a well-functioning restructuring process and that this disruption is closely associated with problems in financial markets. The result is a productivity-based social cost of economic crises that is in addition to the traditional cost from underemployment of labor and underutilization of other resources. The cost of crises in terms of restructuring is twofold: Crises are likely to result in a significant amount of privately inefficient liquidations, leading to large costs in job losses and liquidations of organizational capital. And they are likely to result in a freezing of the restructuring process and years of stagnation in productivity.

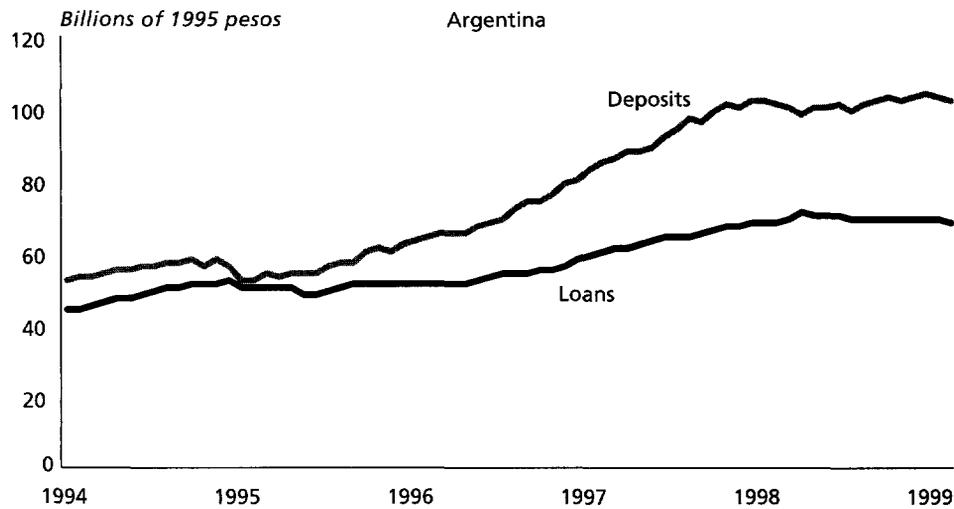
Figure 6. Impulse-Response Functions for a Recessory Shock: A Case of Chill



Note: The regression underlying the figure uses U.S. manufacturing employment (N_t), gross job creation (H_t), and gross job destruction (D_t) in deviation from their mean. The data are quarterly and cover 1972–93. Employment fluctuations are assumed to be driven by a single aggregate shock. Given the identity $DN_t = H_t - D_t$, a linear time-series model for the response of job flows to aggregate shocks can generally be written in terms of either creation, $H_t = qh(L)N_t + e_{ht}$, or destruction, $D_t = qd(L)N_t + e_{dt}$, where $qh(L)$ and $qd(L)$ are polynomials in the lag operator L . The figure shows the estimated impulse-response functions for a two-standard-deviation recessory shock. Source: Caballero and Hammour 1999.

Figure 7. Private Deposits and Loans in Argentina and Mexico, 1990s

Source: Banco de Mexico.



Source: Banco Central de la Republica Argentina.

Conclusion

The massive, ongoing restructuring and factor reallocation by which new technologies replace the old is a core mechanism of economic growth in modern market economies. This process of Schumpeterian creative destruction permeates major aspects of macroeconomic performance—not only long-run growth, but also economic fluctuations and the functioning of factor markets. The process of creative destruction is also fragile, exposed as it is to political short-sightedness, inadequate contractual environments, and financial underdevelopment.

We have reviewed both theoretical arguments and empirical evidence for this creative destruction view of macroeconomic performance. While the evidence we have presented is mostly from developed economies, it is no great leap to conjecture that many of our empirical findings also apply to developing economies. In fact, these economies typically suffer from more severe deficiencies in their contractual environment and from more severe damage to their financial systems during crises—the two most important factors underlying sclerosis as well as inefficient restructuring following contractions.

There is clearly a significant need for new and more structural empirical evidence on the workings of the creative destruction process and its perils in developing economies. We hope that this article has pointed to some of the most promising issues on this agenda.

Notes

1. An alternative empirical approach to creative destruction focuses on physical capital and asks how much of the growth in output is associated with capital-embodied technological progress (see Hulten 1992 and Greenwood, Herkowitz, and Krusell 1997).

2. To be more precise, Aw, Chen, and Roberts use firm- rather than plant-level data and construct a within-firm rather than a within-plant term.

3. Since Aw, Chen, and Roberts (1997) do not provide sector weights, the calculated average contribution gives equal weight to the total factor productivity growth rates in their table 12.

4. One reason for this joint negotiation could be that the entrepreneur can disguise his own funds as coming from external financiers and external funds as being his own.

5. To avoid issues relating to the possibility that the entrepreneur may want to start over in a new production unit, we assume that entrepreneurs in preexisting units have zero net worth.

6. Banerjee and Newman (1998) apply a similar interpretation to the traditional sector, which they see as having easier contracting because information asymmetries are less severe.

7. Note that equation 9 implies that $\pi_i > 0$.

8. See Davis, Haltiwanger, and Schuh (1996) for evidence from U.S. manufacturing. Where analyses have been conducted, they have shown that a large share of destruction during contractions is permanent (see Davis and Haltiwanger 1992).

9. Fluctuations in the pace of restructuring can be approached from a very different angle, by moving from job reallocation to the restructuring of corporate assets. Looking at merger and acquisition activity over time and at its institutional underpinnings, we reach a conclusion that also amounts to a rejection of the liquidationist perspective (see Caballero and Hammour 2000). Essentially, liquidationism in this context would consider fire sales during sharp liquidity contractions as the occasion for intense restructuring of corporate assets. The evidence points, on the contrary, to briskly expansionary periods characterized by high stock market valuations and abundant liquidity as the occasion for intense merger and acquisition activity. Again, financial factors and their institutional underpinnings seem to be at the core of this restructuring phenomenon.

10. It would probably be unwise to look for direct evidence of depressed reallocation along the lines that we use for the United States. The reason is that crises in developing economies often involve large changes in relative prices (such as the large real devaluation during Mexico's tequila crisis), which naturally induce reallocation. The right metric is then one that controls for this purely neoclassical mechanism.

11. The slow recovery of loans in Argentina was caused by the government's crowding out as it borrowed to pay for its monetary intervention and, most important for our argument, by the sharp consolidation in the banking sector following the crisis.

References

- Aw, Bee-Yan, Xiaomin Chen, and Mark Roberts. 1997. "Firm-Level Evidence on Productivity Differentials, Turnover, and Exports in Taiwanese Manufacturing." NBER Working Paper 6235. National Bureau of Economic Research, Cambridge, Mass.
- Baily, Martin, Charles Hulten, and David Campbell. 1992. "Productivity Dynamics in Manufacturing Plants." *Brookings Papers on Economic Activity: Microeconomics*, pp. 187–267.
- Banerjee, Abhijit V., and Andrew Newman. 1998. "Information, the Dual Economy, and Development." *Review of Economic Studies* 65: 631–53.
- Bartelsman, Eric J., and Phoebus J. Dhrymes. 1998. "Productivity Dynamics: U.S. Manufacturing Plants, 1972-1986." *Journal of Productivity Analysis* 9 (1): 5–34.
- Bertola, Giuseppe, and Richard Rogerson. 1997. "Institutions and Labor Reallocation." *European Economic Review* 41: 1147–71.
- Caballero, Ricardo J., and Mohamad L. Hammour. 1996a. "On the Ills of Adjustment." *Journal of Development Economics* 51: 161–92.
- . 1996b. "On the Timing and Efficiency of Creative Destruction." *Quarterly Journal of Economics* 111 (3): 805–52.
- . 1998a. "Jobless Growth: Appropriability, Factor Substitution, and Unemployment." *Carnegie-Rochester Conference Series on Public Policy* 48: 51–94.
- . 1998b. "The Macroeconomics of Specificity." *Journal of Political Economy* 106 (4): 724–67.
- . 1999. "The Cost of Recessions Revisited: A Reverse-Liquidationist View." NBER Working Paper 7355. National Bureau of Economic Research, Cambridge, Mass.
- . 2000. "Technological Revolutions and Merger Waves." DELTA, Paris
- Contini, Bruno, and Lia Pacelli. 1995. "A Study on Job Creation and Destruction in Europe." A Study for the Commission of the European Communities. University of Turin, Department of Economics.
- Cox-Edwards, Alejandra. 1993. "Labor Market Legislation in Latin America and the Caribbean." Regional Studies Program Report 31. World Bank, Latin America and the Caribbean Technical Department, Washington, D.C.
- Davis, Steven J., and John Haltiwanger. 1992. "Gross Job Creation, Gross Job Destruction and Employment Reallocation." *Quarterly Journal of Economics* 107: 819–64.
- . Forthcoming. "Gross Job Flows." In Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*. Amsterdam: North-Holland.
- Davis, Steven J., John Haltiwanger, and Scott Schuh. 1996. *Job Creation and Destruction*. Cambridge, Mass.: MIT Press.
- De Long, J. Bradford. 1990. "'Liquidation' Cycles: Old-Fashioned Real Business Cycle Theory and the Great Depression." NBER Working Paper 3546. National Bureau of Economic Research, Cambridge, Mass.
- Fernandez, Raquel, and Dani Rodrik. 1991. "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty." *American Economic Review* 81 (5): 1146–55.
- Foster, Lucia, John Haltiwanger, and C.J. Krizan. Forthcoming. "Aggregate Productivity Growth: Lessons from Microeconomic Evidence." In Edward Dean, Michael Harper, and Charles Hulten, eds., *New Directions in Productivity Analysis*. Chicago: University of Chicago Press.
- Greenspan, Alan. 1999. "Maintaining Economic Vitality." Millennium Lecture Series, Gerald R. Ford Foundation and Grand Valley State University, 8 September, Grand Rapids, Mich.
- Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell. 1997. "Long-Run Implications of Investment-Specific Technological Change." *American Economic Review* 87 (3): 342–62.

- Hart, Oliver, and John Moore. 1994. "A Theory of Debt Based on the Inalienability of Human Capital." *Quarterly Journal of Economics* 109: 841–79.
- Haltiwanger, John, and Milan Vodopivec. 1997. "Gross Worker and Job Flows in a Transition Economy: An Analysis of Estonia." University of Maryland, Department of Economics, College Park, Md.
- Hulten, Charles R. 1992. "Growth Accounting When Technical Change Is Embodied in Capital." *American Economic Review* 82: 964–80.
- Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica* 50: 649–70.
- Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian. 1978. "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process." *Journal of Law and Economics* 21 (2): 297–326.
- Levinsohn, James. 1999. "Employment Responses to International Liberalization in Chile." *Journal of International Economics* 47 (2): 321–44.
- Lin, Justin Yifu, and Jeffrey B. Nugent. 1995. "Institutions and Economic Development." In Jere R. Behrman and T. N. Srinivasan, eds., *Handbook of Economic Development*, vol. 3A. Amsterdam: North-Holland.
- Liu, Lili, and James R. Tybout. 1996. "Productivity Growth in Chile and Colombia: The Role of Entry, Exit and Learning." In Mark J. Roberts and James R. Tybout, eds., *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*. Oxford: Oxford University Press.
- Melitz, Marc J. 1999. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." University of Michigan, Department of Economics, Ann Arbor, Mich.
- Mokyr, Joel. 1992. "Technological Inertia in Economic History." *Journal of Economic History* 52 (2): 325–38.
- Olley, Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica* 64: 1263–97.
- Ramey, Valerie A., and Matthew D. Shapiro. 1998. "Displaced Capital." NBER Working Paper 6775. National Bureau of Economic Research, Cambridge, Mass.
- Roberts, Mark J. 1996. "Employment Flows and Producer Turnover in Three Developing Countries." In Mark J. Roberts and James R. Tybout, eds., *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*. Oxford: Oxford University Press.
- Roberts, Mark J., and James R. Tybout, eds. 1996. *Industrial Evolution in Developing Countries: Micro Patterns of Turnover, Productivity, and Market Structure*. Oxford: Oxford University Press.
- Schumpeter, Joseph A. 1942. *Capitalism, Socialism, and Democracy*. New York: Harper and Brothers.
- Thomis, Malcolm I. 1972. *The Luddites*. Hamden, Conn.: Archon Books.
- Tybout, James R. 2000. "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?" *Journal of Economic Literature* 38 (March): 11–44.

Comment on “Shaken and Stirred: Explaining Growth Volatility,” by William Easterly, Roumeen Islam, and Joseph E. Stiglitz, and “Creative Destruction and Development: Institutions, Crises, and Restructuring,” by Ricardo J. Caballero and Mohamad L. Hammour

Abhijit V. Banerjee

The titles of these two articles obscure the similarity of their basic messages: that economies that have poor property rights protection and, more generally, weak capital market institutions tend to be more volatile. Ricardo Caballero and Mohamad Hammour provide a set of arguments, based on their important previous work, on why negative shocks (crises) would be expected to have particularly damaging effects in such economies.¹ William Easterly, Roumeen Islam, and Joseph Stiglitz pull together a range of interesting cross-country data to show that there is a strong positive correlation between various indicators of weak capital market institutions and volatility in growth.

A Puzzle

At first glance, the positive correlation between weak capital market institutions and volatility is puzzling. A simple intuitive argument would go as follows: Economies with good capital markets are best able to take advantage of positive productivity shocks (new innovations, terms of trade shocks). Therefore, by symmetry, they will be the ones most damaged by negative productivity shocks. For example, an economy in which capital is very mobile will grow very fast when it receives a positive growth shock, but for exactly that reason its growth will slow the most when the positive growth shock fails to arrive.

My aim is to understand the basis of this correlation and its implications. I begin by briefly examining some potential data problems and some potential explanations.

Explanation 1: Spurious Correlation

Economies with excellent capital market institutions are likely to be richer and, for the usual convergence reasons, may be expected to have a lower marginal product of capital than economies with weak capital market institutions. As a result, shocks that induce marginal firms to shut down may have a smaller effect in economies with good capital

Abhijit V. Banerjee is professor of economics at the Massachusetts Institute of Technology
Annual World Bank Conference on Development Economics 2000
©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

markets. There may therefore be a spurious correlation between the quality of capital markets and growth volatility that results from omitting the marginal product of capital. The results of Easterly, Islam, and Stiglitz do hold even after linearly controlling for per capita GDP, but the authors may not have controlled for enough. Moreover, the effect of a shock on growth volatility may depend on the level of the growth rate (for which they do not control), and economies with better capital markets may grow more slowly for the usual convergence reasons.

This point is worth investigating further, but my sense is that this source of spurious correlation is not especially compelling. The problem is that there is little evidence for unconditional convergence.² Moreover, to the extent that Easterly, Islam, and Stiglitz find any effect of per capita GDP on growth volatility, the measured effect is positive, implying that richer countries are more volatile.

The labor market is another potential source of omitted variables. If labor market rigidity is correlated with capital market failures, measures of the quality of the capital market could be picking up labor market characteristics. Easterly, Islam, and Stiglitz argue against this possibility and, more generally, against the view that labor market rigidities are the source of volatility. Their evidence is that real wage flexibility, as measured by the standard deviation of real wages and the standard deviation of real wage changes, is higher in developing countries and, moreover, almost uncorrelated with growth volatility.

I am less persuaded by this evidence. The problem is that the standard deviation of real wages (or of real wage changes) does not measure labor market flexibility very well. An economy that has very flexible nominal wages may have completely stable real wages if the only source of shocks is the money supply. An economy that has rigid nominal wages but relatively flexible prices (as many formerly socialist economies undergoing liberalization do) will show large variations in real wages. In this case the standard deviation of real wages will reflect rigidity in the labor market rather than flexibility.

Another problem with this measure becomes apparent if we compare an economy in which real wages adjust instantaneously to productivity shocks with one in which there is no adjustment until real wages are far enough from the market equilibrium wages. The first will have many small adjustments, the second occasional large changes. The second, which clearly has the more rigid labor market, could easily have a higher standard deviation of real wages and real wage changes.

I am nevertheless willing to accept their conclusion. The microeconomic evidence from developing countries seems to support the view that these countries have relatively flexible labor markets, at least outside the relatively small formal sector (see, for example, Rosenzweig 1988). Labor market inflexibility is therefore unlikely to be a big part of the story.

Explanation 2: Stabilizing Capital Flows

The conventional explanation of why good capital market institutions have a stabilizing effect comes from the idea that such institutions allow countries better access to world capital markets, enabling them to smooth consumption and investment by

borrowing abroad. If foreign capital helps to reduce volatility in GDP growth, it must be the case that savings drop sharply in a crisis but investment does not. The folk wisdom seems to be that it is investment that typically responds more sharply to a shock in most developing countries, and, in fact, Easterly, Islam, and Stiglitz find no effect of foreign capital flows on volatility. Nevertheless, it may be worth using data on investment and savings volatility outcomes to investigate this explanation.

Explanation 3: Balance Sheet Effects

The basic explanation that Easterly, Islam, and Stiglitz offer for their results relies on balance sheet effects of shocks. Their basic idea, as I understand it, is that firms in economies with weak capital market institutions depend heavily on cash flow to finance production and investment, so that shocks to their balance sheet have strong effects on output and growth.

The problem with this version of the argument is that it does not distinguish sharply enough between levels and changes. It is true that firms in economies with weak capital market institutions finance a larger share of their outlay from internal sources. Put another way, these firms are less able to leverage their internal resources into additional funds than are their counterparts in more developed capital markets. This does imply that the level of output in these economies will be lower. But it does not necessarily imply that their output will be more responsive to changes in cash flow.

There are at least two conflicting effects here. First, a one-unit fall in a firm's cash flow will reduce the total resources available to the firm by more where the capital markets function better (that is, where the firm gets more leverage). The resulting fall in output is therefore likely to be greater where capital markets are better. Second, the better the capital market, the more likely it is that the firm has not already borrowed up to its maximum, which may exceed its needs under normal conditions. This will make the firm's outlays less responsive to short-term changes in cash flow. The net effect is ambiguous, though in some plausible scenarios it can be shown to have an inverted U-shape—that is, an improvement in the capital market first increases volatility and then reduces it (see, for example, Aghion, Banerjee, and Piketty 1999 and Aghion, Bacchetta, and Banerjee 1999).

That the direction of the correlation between the quality of capital markets and growth volatility is unclear does not of course rule out the possibility that it is positive. But it does raise the possibility that the positive correlation is driven by the fact that the data are restricted to a sample of relatively rich countries.³

Explanation 4: Sclerosis and Scrambling

Caballero and Hammour propose another explanation based on the idea that both positive and negative shocks are, in their own way, opportunities. The best way to take advantage of good times is to move capital fast into whatever is hot. The way to benefit from bad times is to move capital out of what has become unproductive in order to make away for what is to come—in other words, creative destruction. In

economies with weak capital market institutions, capital is much too slow both to move in and to move out. This is what Caballero and Hammour call *sclerosis*. Moreover, when capital does move out, it might move in the wrong direction: firms with deep pockets may outlive the most productive firms. In Caballero and Hammour's words, this is a *scrambling* of the natural order of things.

What distinguishes Caballero and Hammour's work from much of the literature on imperfect capital markets is its emphasis on the interaction of such markets with firm-level heterogeneity and the resulting patterns of capital movement. In this view a negative shock becomes an occasion for a certain type of reallocation rather than simply an overall contraction. Countries where such reallocation takes place efficiently can therefore withstand the shock without a substantial drop in their growth rates.

Is the relative immobility of capital resulting from imperfect capital markets large enough to warrant taking Caballero and Hammour's view seriously? While there is no direct evidence, a recent study of the knitted garment industry in Tirupur, in Southern India, provides some indirect support (Banerjee and Munshi 2000).⁴ Tirupur produces 70 percent of the knitted garment exports from India, a major exporter in this category. Tirupur has two communities of producers: Gounders, linked by community ties with a rich local agricultural community, and Outsiders, a motley crew of businessmen from all over India. They produce exactly the same goods, yet use radically different technologies. Gounders invest much more than Outsiders at all levels of experience, both in absolute terms and relative to output. Average capital-output ratios for Gounders can be three times those for Outsiders and are typically twice as large. But all the evidence points to the Outsiders being more able: they enjoy faster output growth, and their output outstrips that of the Gounders after a few years. In other words, the more able invest less in this industry.⁵

The reason, it appears, is that the Gounders have a lot of investible funds that they cannot profitably lend out because capital markets in India function poorly. Instead, they set up garment firms or lend to friends and family in the garment business. Since the firms are set up as a conduit for the surplus capital, they are not required to be particularly productive. The Outsiders, by contrast, come from traditional entrepreneurial communities and their capital probably has many alternative uses. So when they invest in Tirupur, it is not because they lack other choices. This makes them more likely to be productive but also less willing to invest a lot.

What is striking about these results is the extent of scrambling. Gounders can invest almost twice as much and still fall behind in all measures of output. Capital seems to be extremely slow to move to its best users. It is thus possible that a theory like that suggested by Caballero and Hammour could explain the observed correlation between weak capital markets and growth volatility, though we have no direct evidence that it does.

Conclusion

The strong association between weak capital markets and growth volatility in the data needs careful interpretation. While there is good reason to believe that mal-

functioning capital markets have large consequences for growth and stability, little is known about the channel that links them. This is important, as the optimal policy response depends on the exact nature of that channel. In Easterly, Islam, and Stiglitz's world the best policy might be to weaken the link between cash flow and investment. In Caballero and Hammour's world it might be to increase the mobility of capital, perhaps by reducing the costs of firing workers. We clearly need more detailed evidence on how capital markets function in the developing world.

Notes

1. This is not all they do, but given space restrictions I focus on this aspect of their results.
2. Since Easterly, Islam, and Stiglitz do not control for the savings rate and the human capital investment rate in their regressions, the correct comparison is with models of unconditional convergence.
3. This is what would follow if the relationship had an inverted U-shape.
4. The study is based on panel data collected from about 600 garment producers in 1995.
5. This is also consistent with capital and ability being substitutes, but Banerjee and Munshi (2000) show other evidence supporting the more standard view that capital and ability are complements.

References

- Aghion, Philippe, Philippe Bacchetta, and Abhijit Banerjee. 1999. "Capital Markets and the Instability of Open Economies." Massachusetts Institute of Technology, Department of Economics, Cambridge, Mass.
- Aghion, Philippe, Abhijit Banerjee, and Thomas Piketty. 1999. "Dualism and Macroeconomic Volatility." *Quarterly Journal of Economics* 114 (4): 1359–97.
- Banerjee, Abhijit, and Kaivan Munshi. 2000. "Networks, Migration and Investment: Insiders and Outsiders in Tirupur's Production Cluster." Massachusetts Institute of Technology, Department of Economics, Cambridge, Mass.
- Rosenzweig, Mark R. 1988. "Labor Markets in Low-Income Countries." In Hollis Chenery and T. N. Srinivasan, eds., *Handbook of Development Economics*. New York: Elsevier Science Pub. Co.

The East Asian Crisis—Two Years Later

Eisuke Sakakibara

Driven by the information and telecommunications revolution, global capital markets have become extremely volatile and increasingly subject to boom and bust cycles. This article argues that the ideal solution to the instability of a global economy with free movement of capital is to create a genuine lender of last resort. The International Monetary Fund has played the role of lender of last resort in part, but the realities of international politics make it difficult to expand its role significantly. The article proposes ways in which emerging economies, particularly those in Asia, can defend themselves, individually and as a group, against the inherent instability of global capital markets. It argues for strengthening regional cooperation in Asia in such areas as liberalizing trade and investment, developing regional debt markets, coordinating exchange rate policy, and creating a regional financing mechanism.

The East Asian financial crisis prompted wide-ranging analysis of the factors that may have caused it. Early in the crisis discussions focused on macroeconomic fundamentals and structural problems in the countries affected. This was only natural, since the Group of Seven (G-7) and the International Monetary Fund (IMF) had been operating under the premise that sound macroeconomic policies and liberalized markets were the basic requirements for good economic performance. Moreover, since the collapse of the Soviet Union in the early 1990s, policy advisers had emphasized structural reforms to move quickly to an open market economy, along with sound macroeconomic management. This initial reaction to the East Asian crisis was derived from what John Williamson (1990) termed “the Washington consensus.”

Later research revealed, however, that macroeconomic indicators were generally strong among the East Asian countries hit by the crisis. Neither market interest rate spreads nor country risk ratings had indicated macroeconomic weaknesses. There

Eisuke Sakakibara is professor and director of the Global Security Research Center at Keio University, in Tokyo, Japan.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

were structural weaknesses in corporate and national governance, but these had existed for decades. Why they suddenly became a problem in 1997 is difficult to explain. Even the problems of the East Asian financial systems had existed for many years. Moreover, many countries with weak banking systems were not hit by crisis. As Barry Bosworth (1999) rightly points out,

to generate a crisis of the magnitude of East Asia there is a need to link a weak banking system to some other triggering event. . . . In East Asia the trigger was financial liberalization and the effort to link domestic financial markets to those of other countries. Many countries have encountered difficulties in managing this process of financial market reform. (pp. 1–2)

In all the countries hit by crisis in 1994–97, from Mexico to the Republic of Korea, aggregate short-term debt exceeded foreign reserves by a substantial margin (Radelet and Sachs 1998). Gradual recognition of this fact, along with the lack of a lender of last resort in foreign currencies, particularly U.S. dollars, triggered the financial panic, which resembled a domestic bank run. The gap between short-term debt, owed largely to nonresidents, and foreign reserves had resulted in part from liberalization of capital controls.

Monetary authorities attempted to defend the fixed or quasi-fixed exchange rates by intervening in foreign exchange markets. Many economists have argued that this defense of unrealistic exchange rates was one of the main causes of the crisis. The Council on Foreign Relations task force on the future international financial architecture, for example, advised against pegging exchange rates and strongly against using funds from the IMF or G-7 to support “unsustainable pegs” (Institute for International Economics 1999).

True, there were cases, such as Thailand, where the exchange rate was overvalued and market interventions to support the peg depleted foreign reserves, triggering the crisis. Also true, however, is that the adoption of a floating exchange rate regime in the midst of or before the crises in Mexico, Thailand, and Indonesia led to a free fall in their exchange rates, significantly aggravating the situation. One might argue that if these countries had floated their exchange rates much earlier—say, when foreign capital was flowing in—the outcome would have been different. Is there any merit to this argument? Suppose Thailand had floated its exchange rate in 1994–95, when there was investor euphoria about Asia. If Thai authorities had not intervened in the market, the baht would have appreciated more. Of course, if the market were rational, taking into account the exchange rate risk, capital flows to Thailand would have declined. My conjecture is that it would not have happened that way.

For economists, particularly those educated in the neoclassical paradigm, it is natural to assume that market participants will behave rationally. In reality, it turned out to be more profitable to ride with the herd and try to skillfully manage the boom and bust cycle: enjoy the boom but jump ship before the other market participants do. In the global economy, where markets have become so interdependent and rev-

olutionary technological innovations create great uncertainty, the assumption of one stable equilibrium is unwarranted. There are multiple equilibria, and once we leave one equilibrium we are likely to be thrown into great instability. In this kind of situation exchange rate flexibility is not necessarily a good thing. Given investors' euphoric expectations about the future, free-floating exchange rates might have accelerated, not moderated, the boom, resulting in a more serious bursting of the bubble.

This is not to say that exchange rate flexibility is necessarily bad. But under ordinary circumstances price flexibility alone does not solve the boom and bust cycle that characterizes the "virtualized" global market. We need to recognize that boom and bust are sometimes inevitable and that an appropriate mechanism has to be established to minimize the risk that the bust will turn into a systemic market collapse. This leads inevitably to the issue of the interplay between the enormous international capital flows into the liberalized markets of emerging economies—flows that can rapidly reverse direction—and the workings of domestic and international financial systems, including domestic financial authorities and international financial institutions.

Several things would help stabilize global capital markets. Emerging economies should strengthen their banking and financial systems and improve corporate governance. Industrial countries should devise prudent regulations for lending institutions and require greater transparency in their operations. International institutions such as the IMF should improve surveillance. But we know very well that such efforts will not change the fundamental nature of global financial markets, prone to herding, panics, contagion, and boom and bust cycles. World opinion has moved on from the market fundamentalism of the early days of the East Asian crisis to a somewhat more balanced and realistic view.

The finance ministers' report to the G-8 summit in Cologne in June 1999 represents important progress in this respect. Haruhiko Kuroda (1999), Japan's vice minister of finance for international affairs, cites three major improvements recommended by the report to the G-8 summit: an emphasis on orderly sequencing in liberalizing capital accounts, allowing the possibility of short-term capital controls; prudent regulation and disclosure of highly leveraged institutions; and recognition of the need for private sector participation in the IMF rescue plan. As he states, it is quite surprising that the G-7 countries—particularly the United States, which had long endorsed the Washington consensus—agreed to these recommendations, a dramatic shift in their position. The East Asian crisis was so severe, and the contagion so extensive, that even those carrying the flag for market fundamentalism were forced to rethink their positions.

George Soros, the wizard of global financial markets, wrote a book entitled *The Crisis of Global Capitalism* (1998) in which he argues that these global markets are inherently unstable. Expounding on two basic concepts, "reflexivity" and "fallibility," he leads us to the conclusion that "market fundamentalism is today a greater threat to open society than any totalitarian ideology" (p. 16). Many market players may disagree with him, but they at least agree that appropriate public infrastructure,

including tough supervision and oversight by monetary authorities on both the lenders' and the borrowers' side, must complement market discipline.

As the crisis recedes, however, a sense of complacency seems to be spreading, particularly in the private sector. The discussions and recommendations in the G-7, G-20, and other forums may not lead to any fundamental reforms of global capital markets, but end up only in minor interior redecoration. We need to seriously address fundamental issues during this period of calm, because it may well be the calm between two storms. And the next storm could well be bigger than the one we experienced in 1994–99. Instant access to information and the worldwide telecommunications revolution have made capital markets more global and virtual, and thus more unstable and vulnerable to shocks.

Purist Solutions—An International Lender of Last Resort and Capital Controls

Mervyn King (1999), deputy governor of the Bank of England, points to two “purist” or logically clean solutions to the instability of global capital markets: creating an international lender of last resort with free movement of capital and reinstating permanent capital controls. He then concludes that neither is feasible or desirable under today’s international political regime and instead advocates muddling through, or the “middle way.”

The Politics of an International Lender of Last Resort

King discounts the possibility of creating an international lender of last resort by saying “the basic reason is the maxim: ‘it’s the politics, stupid.’” Unfortunately, he is right. Still, it is probably useful, particularly for the countries at the periphery, to analyze the nature of international politics. Since Japan is a little further removed from the center than the United Kingdom is, I may be in a better position than Deputy Governor King to perform this task.

Alan Meltzer (1986) writes that

the central bank is called the lender of last resort because it is capable of lending—and to prevent failure of solvent banks must lend—in periods when no other lender is either capable or willing to lend in sufficient volume to prevent or end a financial panic. (p. 83)

While in the domestic context the central bank acts as lender of last resort, in the current international context neither the IMF nor the IMF and the World Bank combined play this role. However, if all the G-7 countries and the IMF were “willing to lend in sufficient volume to prevent or end a financial panic,” they could serve the function of international lender of last resort. The question is, why couldn’t we—or, more precisely, why were we unwilling to—formalize a lender-of-last-resort mechanism centered on the G-7 during the East Asian crisis?

The answer is clear. Financial panic in one country or region is not necessarily a crisis for other countries. True, there is a possibility of contagion. After the Russian crisis of August 1998 the fear of contagion became real for the United States. But until then the East Asian crisis had not affected the countries at the center of global capital markets. As long as a crisis remains limited to one country or region, unaffected countries face no urgent political need to pay the significant cost associated with playing international lender of last resort. Realism, not altruism, dictates policy in the G-7 and other countries. Moreover, for financial institutions in countries at the center, including mutual and hedge funds, a crisis elsewhere may present an opportunity to increase profits. In fact, except for the brief period from August 1998 to early 1999, U.S. financial institutions and the U.S. economy gained significantly from the East Asian crisis.

Soros (1998) describes the global capitalist system as “purely functional,”

and the function it serves is (not surprisingly) economic: the production, consumption and exchange of goods and services. . . . Despite its non-territorial nature, the system does have a center and periphery. The center is the provider of capital; the periphery is the user of capital: the rules of the game are skewed in favor of the center. (p. 122)

For the countries at the periphery, this bias toward the center is the real issue. Even collectively, they do not have the political leverage to persuade the countries at the center. They could wait for the next crisis—which might really hit the center—for those countries to come around. Or they could adopt a defensive mechanism against recurrent crises, imposing emergency or permanent capital controls or creating what King (1999) calls the do-it-yourself lender of last resort. As defensive measures, both involve forsaking benefits of a free capital market and thus lead to efficiency losses. But given the imperfections of today’s global capital markets and the potential for huge damage, economic and social, from the next crisis, such defensive measures might be the politically correct choice. If it is politics in countries at the center that hamper purist solutions, countries at the periphery have to opt for the politically clever but economically second-best solutions.

Capital Controls—the Best of the Second-Best Solutions?

A good example of a successful defensive measure by a country at the periphery is Malaysia’s imposition of capital controls on 2 September 1998. Economists trained in the neoclassical paradigm have a built-in bias against capital controls, but even those at the IMF had to admit that “controls gave the Malaysian authority some breathing space to address the macroeconomic imbalances and implement banking system reforms” (IMF 2000, p. 13).

The Malaysian National Economic Action Council had prepared a national economic rehabilitation plan in early August 1998, a month before the capital controls were imposed along with exchange controls. Key elements of the plan included sta-

bilizing the value of the ringgit, restoring market credibility, maintaining the stability of the financial market, improving the economic infrastructure, giving priority to social security policies, and rehabilitating every economic sector. The plan aimed to establish a social safety net, increase the transparency of the nation's economic system—while fending off criticism of cronyism—and achieve economic efficiency and a healthy recovery of the financial system.

The plan differed from the orthodox IMF prescriptions in two ways. First, it departed from IMF-style shock therapy, though it took international criticism about cronyism and inadequate transparency seriously. Second, it took a Keynesian approach in its fiscal and monetary policies rather than reflecting the monetarist bias of the IMF recommendations. In the face of strong deflationary pressure from the East Asian financial crisis, adopting Keynesian policies was quite appropriate. Realistic but aggressively implemented structural reforms of the financial system also helped.

As the IMF report (2000) points out, the success of the Malaysian policy also depended crucially on the effectiveness of the controls. That in turn depended on the competence of the central bank, Bank Negara, and the existence of monitoring mechanisms.

The controls were wide-ranging and combined capital controls with exchange controls, but without restricting payments and transfers for current international transactions and foreign direct investment. . . . Practically all legal channels for a possible buildup of ringgit funds offshore were eliminated. Offshore ringgit were required to return onshore, limits were imposed on imports and exports of ringgit currency, the use of ringgit currency in trade payments and offshore trading of ringgit assets were prohibited, and transfers between external accounts of nonresidents and ringgit credit facilities between residents and nonresidents were prohibited. (IMF 2000, p. 12)

With supplementary measures, the ringgit was effectively insulated, and the exchange rate fixed at 3.8 ringgit per U.S. dollar. The IMF noted that

since the introduction of controls, there have been no signs of speculative pressures on the exchange rate despite the marked relaxation of fiscal and monetary policies to support weak economic activity. Nor have there been signs that a parallel or non-deliverable forward market is emerging; and no significant circumvention efforts have been reported. (p. 12)

Not all countries at the periphery have the infrastructure to erect effective capital and exchange controls without risking the spread of corruption. Malaysia did, but it was also helped by the end of the global financial crisis in 1999 and the improving international conditions. Nevertheless, the success of defensive capital and exchange controls in the midst of financial crisis remains. Singapore is another example: it has

insulated its currency for many years. Malaysia and Singapore show that imposing capital and exchange controls does not mean that a country must close its doors completely to the rest of the world. Neither can they be called a closed economy: extensive trade and direct and portfolio investment have taken place in Malaysia and Singapore.

Of course, capital and exchange controls are just one of many defensive policies that small, open economies could use, but the examples here show that in the absence of a true international lender of last resort, countries at the periphery can insulate their economy in selected areas and still reap the benefits of free flows of goods and services. Market fundamentalists often preach that market liberalization is an *all or nothing* undertaking. That is not the case. Countries can, and perhaps should, opt for partial liberalization, depending on their size, stage of development, and social and political environment. In this respect Dani Rodrik (1998) is right in arguing against capital account convertibility:

One wonders which of the ills of international capital markets the proposed medicine (capital account convertibility) will remedy. Will the African countries get the foreign capital they need if they remove capital controls? Will “emerging markets” be less at risk of being flooded with foreign capital when such flows conflict with the domestic goals of inflation control or of maintaining a competitive exchange rate? Will sudden reversals become less likely than before? Will contagion across countries be less severe? Will more of the inflows take the form of long-term physical investments rather than short-term flows? . . . It is not that capital controls are necessarily the answer to these problems; they are not. But capital-account liberalization fits the bill less. (p. 2)

To use King’s term, a defensive “middle-way” approach to capital and exchange controls may be the appropriate second-best solution for many emerging economies.

Developing Regional Capital Markets and Currency Arrangements

In a recent speech Lawrence Summers, secretary of the U.S. Treasury, argued for “a greater focus on the strength of national balance sheets” (1999, p. 3). In view of the experience of the East Asian crisis, I agree with him. Summers (1999) contended that

the IMF should actively promote a more fully integrated assessment of a country’s liquidity and balance sheet. Governments need to think long and hard about their approach to financial liberalization—and, in particular, the danger of opening up to short-term capital in the presence of too many domestic guarantees. And they need to manage the government’s own debt in a way that best insures them against future risks. The most sophisticated debt managers are not those who achieve the lowest possible cost of borrowing. (p. 3)

That the U.S. treasury secretary came around to recognizing the “danger of opening up to short-term capital” is quite interesting. But he is right that management of national debt—or of assets and liabilities, including short-term capital—was a crucial element of the East Asian crisis. In fact, Donald Tsang, Hong Kong’s financial secretary, recognized this in the midst of the crisis, in December 1997, writing that “the Asian currency problem is essentially one of funding mismatch compounded by ineffective intermediations” (1997, p. 68). It was indeed a currency and maturity mismatch that led to the deterioration of the crisis countries’ balance sheets. For Korea and Thailand financial institutions’ short-term borrowings in U.S. dollars were the problem; for Indonesia it was corporations’ short-term U.S. dollar borrowings.

The IMF surveillance of countries’ liquidity and balance sheets suggested by Summers would be useful. But the question is whether it would be possible to compile satisfactory balance sheets for private sector entities, including their offshore and off-balance sheet transactions. If effective capital and exchange controls were in place, these statistics would be collected and scrutinized. It would also theoretically be possible to do so if there were legal reporting requirements backed by penalties. Short of such controls and reporting requirements, however, the role of authorities would be more indirect.

But what are the underlying factors leading to the balance sheet problems? As Tsang (1997) points out,

despite generally strong economic fundamentals, high savings and prudential fiscal policy among its economies, Asia traditionally invests most of its savings outside the region, mainly in OECD markets. Funds flow back to Asia in the form of foreign direct investment and portfolio investment. Indeed, most of Asian’s official foreign reserves are invested overseas in long-term instruments, while coming into Asia are bank loans, direct and portfolio investment, which are largely short-term. (p. 68)

The reason is that Asia lacks deep, resilient debt markets and that the United States and other OECD countries are at the center of global capital markets. Moreover, most Asian investments overseas are in U.S. dollars or, more recently, in euro and yen, and much of the capital that flows back is also denominated in U.S. dollars.

Indeed, during the years preceding the crisis capital inflows to crisis-affected countries far exceeded outflows. Total inflows for five Asian countries—Indonesia, Korea, Malaysia, the Philippines, and Thailand—increased from \$47.4 billion in 1994 to \$92.8 billion in 1996. Foreign reserves increased from \$22.9 billion in 1994 to only \$37.9 billion in 1996. Of the \$92.8 billion in inflows in 1996, \$74.0 billion, mostly short-term capital, went through commercial banks and nonbank private corporations.

Yet in the postcrisis period structural characteristics in East Asia remain such that the region continues to supply substantial liquidity to the world. Two things are imperative for the region: to avoid the dramatic reversals in capital flows that gen-

erate boom and bust cycles and to use its abundant liquidity to smoothly meet capital needs in the region. An obvious solution is to create well-functioning capital markets—or, more specifically, debt markets—in the region.

Creating Asian Debt Markets

During the crisis Tsang advocated using international financial institutions such as the World Bank and Asian Development Bank to kick-start private markets for debt in the region. He argued for having such institutions issue high-quality debt paper in Asian markets to supplement liquidity to crisis countries. The Japanese government did take an initiative intended to boost the development of debt markets. In the second phase of the New Miyazawa Initiative it provided government funds directly and through the Japan Fund of the Asian Development Bank to partially guarantee sovereign debt issued by Asian countries.

The failure to develop well-functioning capital markets has left mobilization of the enormous savings in Asia to domestic and foreign banks, which have not received adequate supervision. Bank managers and supervisors have been given responsibility for monitoring the management of these assets, but in the absence of sufficient market checks the quality of the monitoring has been poor. Maturity transformation and the currency composition of assets and liabilities have been especially problematic. Development of markets for medium- to long-term debt denominated in domestic currencies or in foreign currencies other than the U.S. dollar would have done much to help prevent excessive exposure to short-term U.S. dollar liabilities.

What are the impediments to establishing such markets in Asia? Consider the example of Japan, which probably has the most developed government debt market in the region, though it lags substantially behind London and New York in infrastructure.

In fiscal 1999 Japan eliminated tax barriers, such as the withholding tax and securities transaction tax, at least for Japanese government bonds. But several other issues need attention. Full-scale competitive auctioning needs to be implemented quickly. Only 60 percent of 10-year government bonds, the core product of the Japanese government bond market, are sold through competitive auctioning. The government is just starting to issue a five-year note to serve as a benchmark medium-term note—something that had not been done in the past because of the conflict with the financial debentures of long-term credit banks. More important, the reopening of issues to increase their volume is badly needed to give greater depth to the market.

Quick action is also needed on settlement and clearance, to achieve the real-time settlement critical for global transactions. The Bank of Japan's settlement system needs to be made compatible with the Brussels-based Euroclear system, for example, and a regional clearance system linking Tokyo, Hong Kong, Singapore, Sydney, and other Asian markets needs to be quickly established. Development of repo (repurchase agreement) and futures markets should also be encouraged, in Japan and elsewhere in the region.

The development of bond markets in Japan and other Asian countries has been delayed by banks' traditional dominance of financial intermediation, essentially because bonds are a close substitute for bank lending. Thus creating well-functioning capital markets implies fundamentally restructuring banking operations. Banks should be encouraged to diversify beyond their traditional lending business by increasing their participation in fee businesses and market-oriented operations. Securitization of loans and dealing and trading by banks in government and private securities markets and foreign exchange markets should also be encouraged. Market making by the principal dealers, including banks, is crucial in developing efficient securities and foreign exchange markets.

Many of these comments about capital markets in Japan also apply to those in other Asian countries. All these capital markets—whether in Tokyo, Hong Kong, or Kuala Lumpur—should be closely linked. And they should deal not only in national but also in regional and global issues, with transactions conducted on the basis of a standard legal and financial infrastructure.

Moving toward Regional Currency Cooperation

Closely related to the challenge of creating debt markets is the issue of currency mismatch or excessive dollarization. Recognition of excessive dollarization or rigid pegging to the U.S. dollar, with short-term borrowing in U.S. dollars, as among the main causes of the East Asian crisis has led to contemplation of various alternative exchange rate systems in the region. Current thinking in academia seems to lean toward two extremes: an absolutely fixed regime, such as a currency board or complete dollarization (as in Panama), or free flotation.

These extreme solutions raise concerns. Relatively small, open economies would find it difficult to float their currencies freely, amid vast cross-border capital flows that can suddenly reverse direction, without risking excessive swings in their exchange rates. And while currency boards did not prevent Argentina and Hong Kong from handling the most recent crisis fairly well, Hong Kong paid an enormous price in defending its system. Moreover, since the characteristics of the Hong Kong economy are not necessarily shared by others, its experience with a currency board may not offer lessons for other Asian countries. Excessive rigidity was at the root of the crisis. Thus even such regimes as fixed rates or a currency board system must be managed carefully to allow flexibility when necessary, unless a country or region gives up all sovereignty over its monetary and other policies.

One viable solution might be full or partial currency unification and common or partially shared monetary and other policies. And this may be the direction in which many countries will head in the medium term, with Europe and regions under strong European influence eventually gathering in a euro zone, and the United States and countries with which it has close ties binding together in a U.S. dollar zone.

What about Japan and the other countries of Asia? They would, of course, have the option of eventually participating in either the euro or the U.S. dollar zone. Another option would be to develop a third currency zone, in Asia. Given the diver-

sity among Asian countries, however—in race, culture, history, and stage of development—achieving unification like that in the European model would be extremely difficult. Nor would it be possible for Japan to play the kind of role that the United States might play in the Americas.

A possible alternative would be to develop regional cooperation in trade, investment, and exchange rate systems. Such cooperation might involve forming a basket of regional currencies—an Asian currency unit—and attempting to use this unit as the denominator for trade and exchange transactions. The nations of the region could then eventually develop a scheme like that based on the European currency unit (ecu).

An Asian Monetary Fund

A strategy of creating regional defensive mechanisms—regional debt markets and regional currency cooperation—would not deny the global “middle way” that King recommends. Nor would it block efforts by Summers and others to bring the IMF closer to being a genuine international lender of last resort by narrowing its functions to meeting the liquidity needs of countries affected by crisis. But it recognizes that the asymmetry in today’s global financial system means that countries at the center are less likely to devote resources to dealing with crises that remain regional—and that the current calm may strengthen their complacency.

- A potentially important regional defensive policy would be to create what King (1999) calls a do-it-yourself lender of last resort, with the aim of providing self-insurance against a liquidity crisis. King suggests several ways of providing such insurance:
- Build up large foreign currency reserves. China already has substantial foreign exchange reserves (\$147 billion at the end of June 1999, and Korea raised its reserves from a low of \$7.3 billion in November 1997 to \$64.8 billion in August 1999. This is not an efficient use of scarce capital, but may be necessary in the absence of more efficient solutions to the risk of crisis. Building up net reserves—through current account surpluses—would reduce world demand at a time that the U.S. economy is likely to provide a smaller stimulus than during the second half of the 1990s. Creating gross reserves by borrowing from abroad and investing the proceeds in liquid international securities also involves costs.
- Set up contingent credit facilities with international banks, as Argentina has done with its contingent repo facility, or collateralized loan facilities along the lines suggested by Martin Feldstein (1999).
- Create regional self-insurance funds.

The recent financial crises make it likely that all these approaches will be pursued. In August–September 1997, for example, the Japanese government proposed a variant of a regional fund, to be known as the Asian Monetary Fund. The idea was essentially to pool part of the foreign reserves of countries in the region. If China, Korea, Japan, and other East Asian countries provided, say, half their reserves to the

fund with specific arrangements for its activation, the fund could serve as an effective regional lender of last resort during the next liquidity crisis.

The Asian Monetary Fund proposal was strongly opposed by the United States and European countries, however, on the grounds that it would undermine the discipline imposed by the IMF and could pose a serious moral hazard problem. But if the fund's function were very narrowly defined as meeting liquidity needs during a crisis, with a specific formula for private sector participation (along the lines of the Korean model, for example), it could complement the IMF's current function. True, if Summers's proposal for restructuring the IMF were implemented, significant overlap might occur. But even then the Asian Monetary Fund's role could be restricted to providing liquidity, with conditions for private sector participation, while the IMF provided surveillance and macroeconomic policy recommendations.

The amount of liquidity that countries would wish to provide to deal with a crisis may differ depending on their risk of contagion. It is only natural for countries close to a crisis that could easily spread to try to contain it by providing liquidity. Indeed, this could be done through bilateral aid, as Japan did under the New Miyazawa Initiative after the Asian Monetary Fund proposal was shot down. Such bilateral aid should probably be formalized as a permanent regional mechanism, however.

Although moral hazard is a serious problem, the moral hazard argument should not be used to distract us from the need for an international lender of last resort with free movement of international capital. The mere existence of a central bank does not pose any moral hazard problem. It is the conditions under which the bank provides liquidity during a crisis that may give rise to moral hazard. The general rule à la Bagehot (1873) of lending with good collateral at punitive rates could be modified, adding a general scheme for private sector participation that leaves substantial discretion for dealing with individual cases. Agreement on lending rules may be easier to reach regionally than globally, given similarities among countries within a region.

Some economists have used the moral hazard argument to defend market fundamentalism. If markets were perfect, maintaining a stable balance between demand and supply, there would be no need for central banks. But once we agree on the need for central banks in the domestic context, the need for an international lender of last resort with free movement of capital cannot be denied. The issue should be what conditions the lender of last resort would impose, not whether the infusion of public funds should be reduced or eliminated. During a liquidity crisis the amount of public funds infused could be very large even with appropriate private sector participation. Countries unaffected by a crisis should not use the moral hazard argument to avoid responsibility. If they have no political incentive to contribute to a fund in their region, they should simply say so.

A regional fund serving as lender of last resort could be combined with other regional cooperation, such as a regional arrangement for exchange rates—a regional currency unit and a mechanism for defending exchange rates within a certain range. Of course, the fund should not be used to defend unrealistic exchange rates. But the common fund would make joint intervention more plausible. This mechanism would

require a common exchange rate policy in the region—cooperation that would have other advantages, given the heavy interdependence among the region's countries.

As in Europe, monetary or international financial cooperation should be accompanied by cooperation in the real sector. An agreement on free trade or on direct investment should probably precede, or at least be pursued simultaneously with, cooperation on a fund and on an exchange rate arrangement. If cooperation in trade and direct investment proceeds with the creation of a regional debt market, a common fund with exchange rate cooperation might develop into an Asian currency zone independent of the euro and U.S. dollar zones. I remain hesitant to recommend aggressively pursuing an Asian currency zone in the short term, since to do so promises to be an enormous challenge. But in the absence of an international lender of last resort, it is an option worth debating within the region and beyond it as we enter the 21st century.

References

- Bosworth, Barry. 1999. "Asian Miracle, Crisis and Recovery." Paper presented at Asian Development Bank Institute workshop on development paradigms, 10 December.
- Feldstein, Martin. 1999. "A Self-Help Guide for Emerging Markets." *Foreign Affairs* 81 (March–April): 93–109.
- G-7 Finance Ministers. 1999. "G-7 Finance Ministers on Global Financial Structure." Paper presented at the G-8 summit, Cologne, 18–20 June.
- IMF (International Monetary Fund). 2000. "Country Experiences with the Use and Liberalization of Capital Controls." Staff Report. Washington, D.C.
- Institute for International Economics. 1999. "Safeguarding Prosperity in a Global Financial System: The Future International Financial Architecture." Report of an independent task force, sponsored by the Council on Foreign Relations. Washington, D.C.
- King, Mervyn. 1999. "Reforming the International Financial System: The Middle Way." Paper presented at a session of the Money Marketeters of New York University, Federal Reserve Bank of New York, 9 September, New York.
- Kuroda, Haruhiko. 1999. "The Third Way: International Financial System." *Yomiuri Shinbun*, 18–19 November, Tokyo. (In Japanese.)
- Meltzer, Alan. 1986. "Financial Failures and Financial Policies." In G.G. Kaufman and R.C. Kormendia, eds., *Deregulating Financial Services: Public Policy Flux*. Cambridge, Mass.: Ballinger.
- Radelet, Steven, and Jeffrey D. Sachs. 1998. "The East Asian Financial Crisis: Diagnosis, Remedies, Prospects." *Brookings Papers on Economic Activity* 1: 1–74.
- Rodrik, Dani. 1998. "Who Needs Capital-Account Convertibility?" *Essays in International Finance*. Princeton, N.J.: Princeton University.
- Soros, George. 1998. *The Crisis of Global Capitalism*. New York: Public Affairs.
- Summers, Lawrence. 1999. "The Right Kind of IMF for a Stable Global Financial System." Paper presented at London School of Business, 14 December, London.
- Tsang, Donald. 1997. "The Asian Debt Market." *Asiaweek*, 19 December.
- 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.

Comment on “The East Asian Crisis—Two Years Later,” by Eisuke Sakakibara

Andrew Berg

Two themes run through the article by Eisuke Sakakibara. One is that crises in emerging markets, both the recent Asian crises and, implicitly, those we may have to worry about in the future, are driven not by fundamentals but by “hot money” and irrational “jumps” between different equilibria. While Sakakibara gives an occasional nod to weak banking systems and other structural problems, these are largely a sideshow.

The second theme is the importance of regional—in particular, Asian—responses to the problem of how to manage international capital markets.

The Pure Liquidity Response

Sakakibara’s view is that liquidity shocks, almost alone, caused the Asian crises, so that a pure liquidity response would have been appropriate. But it would be a mistake to conclude that all the analysis over the past couple years has led to a consensus around this view. Just as bank runs in the United States before the creation of the Federal Reserve System in 1914 often came not out of the blue but in response to perceptions of real weaknesses, so real weaknesses were a key part of the problem in the Asian crisis countries.

These countries showed deep problems before the crisis, not just inadequate reserves and excessive short-term external debt. Thailand, of course, along with Malaysia, showed important signs of traditional macroeconomic problems, particularly exchange rate overvaluation and excessive current account deficits. In work with Catherine Pattillo, I have shown that problems with traditional macroeconomic fundamentals in these countries led to a high risk of crisis (Pattillo and Berg 1999). In contrast, Indonesia and the Republic of Korea presented few signs of tra-

Andrew Berg is deputy division chief of the Developing Countries Studies Division at the International Monetary Fund. At the time this paper was written he was deputy assistant secretary for East Asia and Latin America at the U.S. Department of the Treasury. The interpretations expressed here are entirely those of the author and do not necessarily represent the views of any organization with which he is or has been affiliated.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

ditional macroeconomic vulnerabilities. But a range of studies of the precrisis situation, from the study by World Bank economists Stijn Claessens and Tom Glaessner (1997) to the prescient book by Adam Schwarz (2000) on Indonesia, outlined the deep structural weaknesses in these economies.

Sakakibara recognizes these issues. But he takes a couple of positions that cast doubt on their importance: First, while plenty of countries have structural weaknesses of this kind, not all of them have crises. Second, these weaknesses had been around for years or even decades, so they cannot explain the timing, at least, of the crises.

The obvious resolution of these two points of view is that structural weaknesses set the stage, but that the timing was a function of external shocks, contagion, and other factors. But against the argument that illiquidity was the only key consideration, consider the case of Malaysia. It is not quite right to say, as Sakakibara does, that while plenty of countries had weak banking systems, only the currencies of the illiquid countries were attacked. Malaysia, with short-term external debt at less than half its reserves, did not have a serious external liability problem. Yet it suffered a crisis about as severe as Korea's and only a bit less so than Thailand's (with severity measured by decline in real output). Moreover, in many of these countries, notably Korea, underlying problems were intensifying.

It is surely true, nonetheless, that illiquidity—measured, for example, by a high ratio of short-term external debt to reserves—played an important part in the incidence and severity of the crises and that stability and growth could not be restored until substantial foreign liquidity was provided. But should providing liquidity have been the entire response?

The pure liquidity response—to what was then an incipient crisis—is exactly what Thailand attempted through late 1996 and early 1997 and, indeed, what Korea tried until late 1997. It didn't work. As these cases show, a country can easily run through huge stocks of reserves defending its exchange rate or accommodating capital outflows through sterilized intervention—that is, by selling reserves (or using other, less transparent mechanisms) while preventing interest rates from increasing significantly.

An adequate infusion of liquidity would presumably have bought substantial breathing room. But the amount of foreign exchange that would have had to be available would have dwarfed the already huge financial support programs. It might be argued that availability alone would have been enough to stem the outflows. But while there is surely some truth to this argument, experience in Mexico (in 1995) and elsewhere suggests strongly that merely promising money is not enough; there are always uncertainties about its true availability. Substantial disbursements are also required, as well as substantial progress in economic reforms.

The Need for Conditionality

Sakakibara spends little time on the role of macroeconomic and structural conditionality, a well-worn topic. But he makes two interesting points with respect to

Malaysia. First, he emphasizes the importance of financial system reforms and efforts to improve transparency and reduce corruption in turning things around in Malaysia. It is reasonable to suspect that in the other crisis countries the large official financing packages reduced the need and therefore the incentive to engage in this kind of difficult reform. This implies a role for conditionality.

Second, Sakakibara lauds Malaysia's lack of macroeconomic austerity. Indeed, Malaysia's monetary policy response was somewhat weaker than those in countries with International Monetary Fund (IMF) programs. While the depreciation of the ringgit was relatively large, reflecting this fact, Malaysia did not therefore have a much smaller recession than the other economies. Indeed, given Malaysia's apparently somewhat stronger initial conditions—in particular, the absence of major external vulnerability—the size of its recession is surprising.

One need not be obsessed with the problem of “moral hazard” to worry about disbursing large loans with no conditions to economies with a deeply troubled financial system—in particular, profound problems relating to the government's role in the system. Perhaps the immediate problems could have been avoided with huge increases in the financing packages available and no structural or macroeconomic reforms. But the adjustments that the recipient countries would eventually have had to make would have been that much larger.

While large, unconditional liquidity support may have theoretical merits, there are also practical limits. The experience in Mexico and Asia showed that political accountability requires that assistance be conditional. The Mexican experience also gave many people a greater appreciation for the importance of having the crisis country and a third party, such as the IMF, work out the conditions.

Capital Controls and Currency Arrangements

Given the roots of crises in external factors and liquidity problems, what are the possible responses? Sakakibara proposes various kinds of solutions to the problem of unstable global capital markets. Here I focus on just a few.

One solution that he proposes is capital controls, but I fear that he fails to come to grips with the issues. Although he makes much of the Malaysian experiment with capital controls since the crisis, there is a much broader set of experiences with capital controls of various kinds, and these have produced several lessons.

First, capital controls cannot prevent large outflows that would normally occur, particularly of capital held by residents, from countries with fairly well-developed financial markets. Second, while controls on short-term capital inflows may at times be justified as prudential measures, they are by no means a panacea: Thailand enacted such controls in the mid-1990s. Moderate controls, such as the famous Chilean-style tax on short-term capital inflows, have small effects at best. Finally, while financial liberalization can be dangerous and is easy to get wrong (such as liberalizing short-term bank lending before foreign direct investment, as in Korea, or giving favorable treatment to short-term external liabilities, as in the Philippines and Thailand), there is little serious debate about the

first-best goal: a modern, open financial market, both domestically and, ultimately, internationally. Even China and Malaysia, for example, are firmly on the path toward this goal.

Nothing in Malaysia's limited experience in the past few years contradicts any of these conclusions. Malaysia's controls on outflows went untested. They were imposed only after the exchange rate had been largely stabilized, with broadly appropriate supportive policies, and during a time that most currencies in the region were tending to appreciate. That no parallel markets emerged suggests not that the controls were remarkably effective, as Sakakibara implies, but that they were not on the whole binding, having been imposed when the worst of the crisis was over and appropriate fundamental reforms were under way.

The upshot of these conclusions is that capital controls are not a real long-term solution to global financial instability. Liberalization is risky and should be pursued carefully, but for countries that have already gone far in this direction, as the Asian crisis countries have, capital controls are no solution. Measures to reduce the risks associated with short-term flows are reasonable but not likely to make a big difference, and adoption of a Chinese-style system of comprehensive capital controls is inconsistent with long-term goals of developing modern financial markets.

Sakakibara also proposes regional currency arrangements. He suggests that while a regional currency union is a long way off for Asia, the region might want to work toward an arrangement of regionally pegged currencies—similar to the precursor to the European Monetary Union—backed by the sharing of reserves or regional swap arrangements. Second thoughts are in order here. The European system proved extremely unstable initially, of course, and its experience suggests that only the near-term goal of a common currency can "anchor" such an arrangement. Moreover, in Europe countries followed Germany's lead in monetary policy. In the absence of such a clear leader, a tremendous degree of political cooperation and discipline would likely be required.

Conclusion

In contrast with a few years ago, most countries in Asia are heading toward a similar long-term development model. While the application of this model will certainly have "Asian characteristics," these characteristics will be less pronounced than in earlier models: Asian countries now generally recognize the need for deeper financial markets, including a greater role for bond markets, and for regulation and supervision in place of the more intimate and multifaceted relations between the state and the financial sector. And rule of law has a stronger place.

How do we avoid crises, and how do we deal with those that do occur? Most of those debating the issue, including Professor Sakakibara, share a consensus on what is needed: more attention to national balance sheets, less dependence on short-term capital flows, more transparency, stronger banking systems, and so on. Do capital controls have a much bigger role to play? Probably not. Will there be a regional currency regime in Asia backed by a regional monetary fund or central bank? Probably not for decades.

History suggests that crises cannot be avoided entirely. Clearly, however, we need to do more to avoid them and to mitigate their impact. Unfortunately, much of what is necessary may not be exciting new initiatives but more determined attention to boring old lessons.

References

- Berg, Andrew, and Catherine Pattillo. 1999. "What Caused the Asian Crises? An Early Warning System Approach." *Economic Notes* 28 (3): 285–334.
- Claessens, Stijn, and Tom Glaessner. 1997. *Are Financial Sector Weaknesses Undermining the East Asian Miracle?* Washington, D.C.: World Bank.
- Schwarz, Adam. 2000. *A Nation in Waiting: Indonesia's Search for Stability*. Boulder, Colo.: Westview.

Mexico—Five Years after the Crisis

Daniel Lederman, Ana María Menéndez, Guillermo Perry, and Joseph E. Stiglitz

This article investigates the main factors in the recovery of the Mexican economy after the currency crisis of 1995. It finds that the V-shaped behavior of fixed investment growth is most important in explaining both the economic slowdown during the crisis and the strength of the subsequent recovery. Econometric results show that fixed investment growth fell precipitously in 1995 as a result of the sharp depreciation of the currency—both because of the negative income effect and because of the increase in economic volatility. But after this initial contractionary effect of the depreciation a substitution effect dominated that favored tradable output, which in Mexico has a larger multiplier effect on investment than does nontradable output. One policy implication that emerges from the analysis is that for countries like Mexico currency devaluation, despite its initial contractionary effect, appears to be a reasonable policy response to speculative attacks.

The Mexican economy contracted sharply after the peso devaluation of 22 December 1994, with GDP growth slowing to an annualized rate of -6.2 percent in 1995. Growth became positive again in the first quarter of 1996, and the economy grew at a healthy rate of 5.1 percent that year and 6.8 percent in 1997. Some authors have suggested that the decisive response of fiscal and monetary authorities, supported by the generous financial package announced by Mexican authorities on 9 March 1995, played a crucial role in the rapid recovery, restoring investor confidence, quickly stabilizing the currency, and turning investment and economic activity around.¹ And some claim that the lack of equally forceful responses and equally large financial support packages might explain why a quick recovery from the recent crises in Asia was more elusive.

Daniel Lederman and Ana María Menéndez are economists in the Office of the Chief Economist for Latin America and the Caribbean at the World Bank. Guillermo Perry is chief economist for Latin America and the Caribbean at the World Bank. Joseph E. Stiglitz is professor of economics at Stanford University. The authors are grateful to Luis Servén, Eliana Cardoso, William Maloney, Alejandro Werner, and two anonymous reviewers for insightful comments on an earlier version of this article. All remaining errors may be attributed to the authors.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

This assessment of the factors behind Mexico's rapid recovery in 1996–97 after the "tequila crisis" of 1995 aims to contribute to the current debate about appropriate responses to currency crises. We use a simple decomposition analysis to establish the contributions of components of aggregate demand to GDP growth since 1993. We examine the elements behind the increase in export growth during and after the tequila crisis and, using a standard model of investment behavior, the determinants of the growth of fixed investment. Finally, we look at the dynamic relationship between fixed investment and the real exchange rate, and investigate whether real exchange rate variations had short-term substitution effects on the composition of output growth.

Decomposition of Growth

A simple decomposition analysis shows that economic growth in Mexico recovered after the tequila crisis primarily as a result of the behavior of fixed investment. The fixed investment growth rate declined drastically during 1995, the year of the crisis, but subsequently rose above precrisis levels. The growth of exports (in constant pesos) seems to have prevented an even deeper recession in 1995, but its contribution to economic growth declined during the recovery.

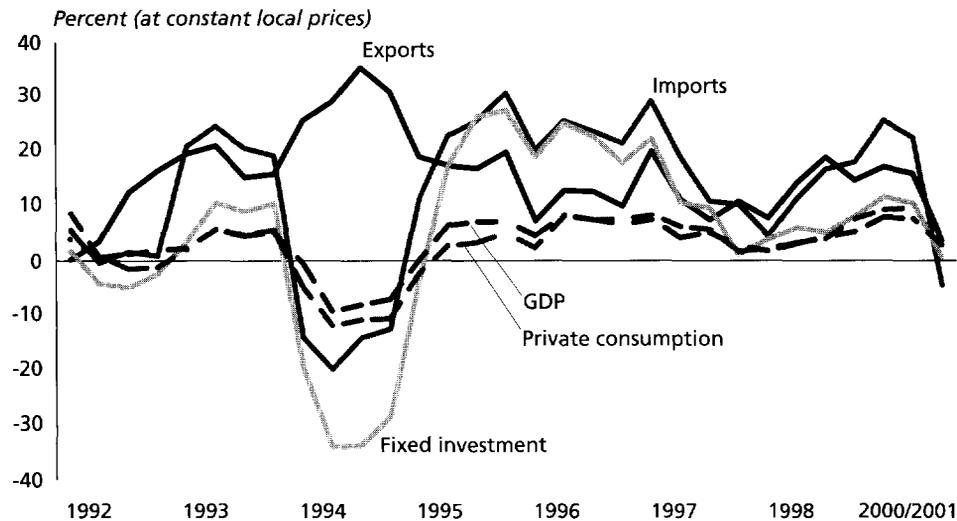
The decomposition of GDP growth into the contributions of the components of aggregate demand comes from a transformation of the basic macroeconomic identity $Y = C + I + X - M$, where Y is output, C is the sum of private and public consumption, I is gross investment, X stands for exports, and M for imports. Calculating first differences for each element, dividing by the ex ante level of Y , and performing some simple manipulations makes it possible to decompose the growth of output into the contributions of the growth of each of its components:

$$(1) \quad \Delta Y/Y_{t-1} = (C_{t-1}/Y_{t-1})(\Delta C/C_{t-1}) + (I_{t-1}/Y_{t-1})(\Delta I/I_{t-1}) + (X_{t-1}/Y_{t-1})(\Delta X/X_{t-1}) - (M_{t-1}/Y_{t-1})(\Delta M/M_{t-1}) .$$

This expression can be easily expanded to include the contribution of more disaggregated components of demand. For example, the contribution of gross investment growth can be further decomposed into the sum of the contributions of fixed investment growth and inventory accumulation.

In the precrisis period exports and imports were the most dynamic components of GDP (figure 1). Export growth rose gradually from a very low rate of 0.18 percent in the first quarter of 1993 to 20.97 percent in the second quarter of 1994. Import growth accelerated more dramatically, from 0.92 percent in the last quarter of 1993 to 24.60 percent in the second quarter of 1994. This sharp rise in the import growth rate reflects the effect of the real exchange rate appreciation that occurred before the devaluation at the end of 1994.

The tequila crisis was sparked by the announcement of a 15 percent nominal devaluation of the peso in December 1994. By March 1995 there was a 43 percent depreciation of the exchange rate. Between the last quarter of 1994 and the second quarter

Figure 1. Average Quarterly Growth of Components of Demand in Mexico, 1992–2000

Source: Authors' calculations; National Institute of Statistics, Geography, and Information (INEGI), Mexico.

of 1995 there was a rapid reversal in the import growth rate, from 19.17 percent to -19.81 percent. The import growth rate remained negative throughout 1995, even though liberal trade policies remained in place. In contrast, exports continued to grow.² These external sector responses to the crisis helped slow the decline of GDP.

By the first quarter of 1996, a year after the crisis began, Mexico's GDP growth rate once again became positive. During the recovery export growth slowed even though exports remained among the most dynamic components of demand. Between 1996 and the third quarter of 1999 export growth averaged 12.99 percent. Meanwhile, imports grew rapidly—by 18.43 percent—surpassing the average growth rates of all other components of demand.

Fixed investment was also dynamic—before, during, and after the tequila crisis. The behavior of fixed investment growth paralleled that of import growth, perhaps because Mexico imports more than 30 percent of its capital goods. During 1994, the year before the crisis, fixed investment growth accelerated from 3.50 percent in the first quarter to 10.43 percent in the last quarter. In 1995, during the crisis, the growth rate of fixed investment dropped more than that of any other component of demand, plummeting to around -30 percent. But fixed investment recovered quickly during the economic recovery, surpassing its precrisis growth rates. Between 1996 and the third quarter of 1999 its growth averaged 14.1 percent.

During the crisis there was a reversal in the contribution of imports to GDP growth. While in 1994 the rapid growth of imports slowed GDP growth to 4.45 percent (reducing it by 4.08 percentage points), in 1995 the slowdown in import growth helped limit the decline in GDP to 6.18 percent (reducing it by 3.35 per-

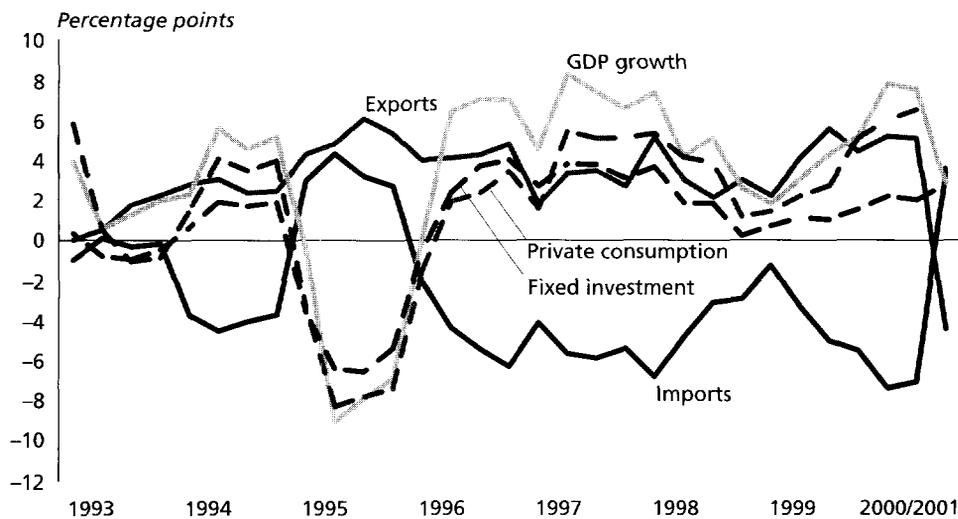
centage points; figure 2). During the recovery import growth rose, once again putting a brake on GDP growth. The contribution of export growth to GDP growth increased remarkably during the crisis, from 2.71 percentage points in 1994 (with GDP growth of 4.45 percent) to 5.19 percentage points in 1995 (with GDP growth of -6.18 percent). During the recovery the role of exports declined: between the first quarter of 1996 and the third quarter of 1999 export growth accounted for 3.42 percentage points of the average GDP growth of 5.13 percent.

In contrast, the contribution of fixed investment fell along with GDP growth in 1995. In that year the decline in fixed investment explained most (5.58 percent) of the 6.18 percent decline in GDP. By 1996 fixed investment had recovered sharply: in the fourth quarter of that year fixed investment growth explained about 60 percent of the GDP growth of 7.11 percent (4.07 percent). Fixed investment growth continued to be important during the postcrisis period, accounting for, on average, 2.27 percentage points of GDP growth, which averaged 5.13 percent.

Export Growth

The external sector has been a key element of the development policies and strategies of the past three Mexican governments. In 1985 the country began a process of trade liberalization by joining the General Agreement on Tariffs and Trade (GATT), the predecessor of the World Trade Organization. In 1991 Mexico entered into negotiations with Canada and the United States over the terms of the North American Free Trade

Figure 2. Average Quarterly Contribution of Growth in Components of Demand to GDP Growth in Mexico, 1992–2000



Note: GDP growth is given as the average quarterly percent change.

Source: Authors' calculations; National Institute of Statistics, Geography, and Information (INEGI), Mexico.

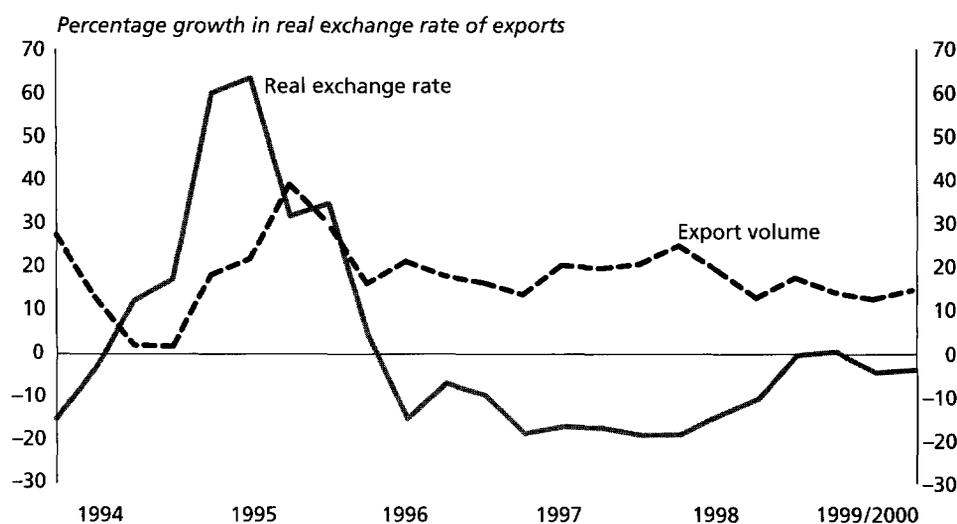
Agreement (NAFTA), which was officially implemented on 1 January 1994. By the onset of the tequila crisis most of the resulting trade reforms had been implemented, and none was reversed as a result of the crisis.³ In this section we look at the role of export growth by analyzing changes in the volume of non-oil exports, measures of competitiveness, and the real exchange rate of Mexican exports (defined as the ratio of the unit price of exports in pesos divided by the Mexican consumer price index).

As figure 1 shows, the growth of exports (at constant local prices) increased during the tequila crisis, while the growth of imports declined—both tendencies that helped to cushion the fall in GDP in 1995. But this peak in the contribution of net exports to GDP growth was short-lived (see figure 2). After the crisis the contribution of imports diminished substantially, and that of exports also declined.

The rise in export revenue growth during the tequila crisis was not all due to changes in the relative price of exports valued in domestic currency, which could be considered a pure “accounting” effect.⁴ It was also due to an increase in the growth of the volume of non-oil exports, which accelerated from 2 percent in the fourth quarter of 1994 to almost 40 percent in 1995 (figure 3). The growth in export volume slowed by the first quarter of 1996, when the economic recovery began, as the real exchange rate of exports began to appreciate. Nevertheless, export volume growth continued to be high, averaging 17.36 percent from the first quarter of 1996 to the third quarter of 1999.

What explains the increase in the growth of export volume? In 1993–99 unit labor costs in Mexican manufacturing fell steadily; before the crisis this occurred even as the real exchange rate was appreciating (figure 4). But the manufacturing remunera-

Figure 3. Average Quarterly Growth in Volume of Non-Oil Exports and Real Exchange Rate of Exports from Mexico, 1994–99



Note: Positive variations of the real exchange rate = depreciations.

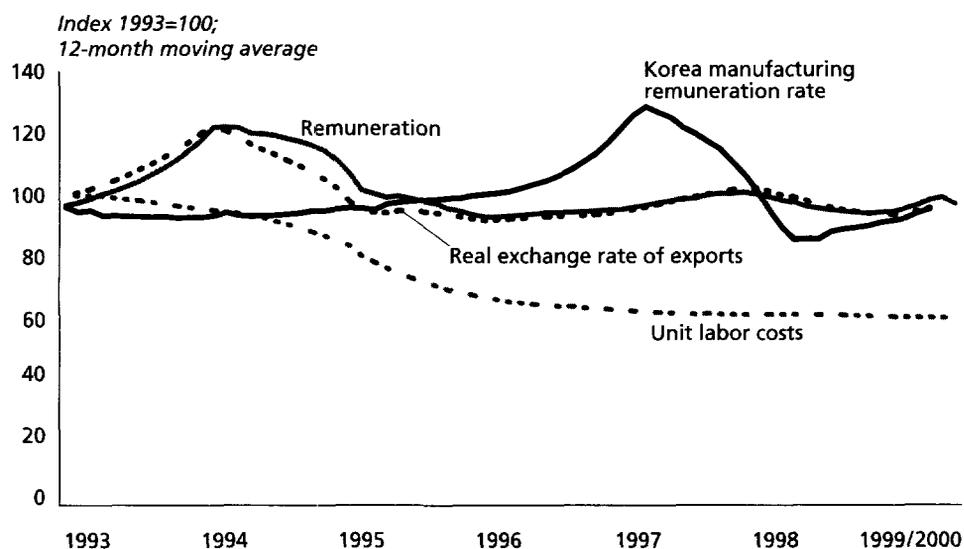
Source: Authors' calculations; National Institute of Statistics, Geography, and Information (INEGI), Mexico.

tion rate in U.S. dollars rose before the crisis. This represented a loss in competitiveness for Mexican exports, associated with the appreciation of the real exchange rate of exports in the same period. In 1995, the year of the crisis, the remuneration rate fell 20 percent as the real exchange rate depreciated. In 1996 both indicators of international competitiveness returned to levels recorded in 1993. We can infer that the gain in competitiveness reflected in the depreciation of the real exchange rate and the decline of the remuneration rate was an important factor in the growth of the volume of non-oil exports in 1995.

In Korea during 1997, before the crisis, the remuneration rate in manufacturing in U.S. dollars increased by 30 percent. It fell through 1998, returning to its precrisis level by the end of 1999. Thus the Mexican and Korean remuneration rates behaved in remarkably similar ways—increasing before the onset of the crisis and returning to their precrisis level the year after the crisis. According to the remuneration indicator, Mexican exports lost competitiveness relative to Korean exports the year before and regained competitiveness the year after the tequila crisis. In 1996 the two countries had similar remuneration rates. By the end of that year Korean exports began to lose competitiveness relative to Mexican exports. By 1999 the countries once again had similar remuneration rates.

In sum, the increase in the growth of Mexican exports observed during and after the tequila crisis was probably driven by an increase in the volume of non-oil

Figure 4. Remuneration in U.S. Dollars and Unit Labor Costs in Manufacturing in Mexico, 1994–99



Note: Data refer to Mexico unless otherwise specified.

Source: Authors' calculations. For remuneration and unit labor costs, National Institute of Statistics, Geography, and Information (INEGI), Mexico. For real exchange rates, International Monetary Fund, various years, *International Financial Statistics*.

exports, which in turn was stimulated by a sustained reduction in unit labor costs and a transitory improvement (or correction) in the international competitiveness of Mexico's industrial sector. That increased competitiveness seems to have been associated with the real depreciation of the peso.

A Mexican Investment Function

The growth decomposition exercise indicates that fixed capital formation was at the heart of the economic recovery in Mexico after 1995. Evidence also suggests that investment in Mexico may be linked to the performance of the tradable sector (Krueger and Tornell 1999). Thus it is important to study more carefully the determinants of investment and its links with the tradable sector.

We can model Mexican investment as a function of the output of the tradable and nontradable sectors (table 1 lists the variables and their sources), where we expect to find different sector multiplier effects:⁵

$$(2) \quad \frac{I}{GDP} = \bar{I} + Y_T^{m_T} + Y_{NT}^{m_{NT}} + I(RIR, p_K, \sigma_{RER})$$

where I/GDP is the level of private fixed investment to GDP ratio at constant prices, \bar{I} is a constant, minimum level of private fixed investment, and m is the multiplier of the corresponding sector: tradable (T) and nontradable (NT). The last term on the right-hand side of the equation is the portion of the investment function that is determined by "cost factors," including the real interest rate (RIR), the relative price of capital goods (p_K), and an uncertainty variable that we identify as the volatility of the real exchange rate (σ_{RER}). This simple model is broadly consistent with standard empirical models of investment behavior in developing countries (see Rama 1993), except for the assumption that tradable and nontradable output growth have different multiplier effects on investment.

Servén (1998) finds a strong negative effect of real exchange rate uncertainty on investment-output ratios in a cross-country panel framework. This effect could be important in explaining the fall in investment in Mexico during 1995 and even during the debt crisis of 1982–83. Moreover, the consideration of real exchange rate uncertainty is consistent with two plausible assumptions. First, domestic investors can be risk averse, so uncertainty may adversely affect private investment. Second, it is also possible that at least portions of private investment are irreversible and contribute to sunk costs. Under these circumstances macroeconomic uncertainty can be associated with swings in the value of private firms, hampering firms' productive investment (Pindyck 1988). If the tradable sector has a large multiplier effect on private investment, real exchange rate uncertainty could also have an indirect effect on investment through its effect on the output of tradables. As Maloney and Acevedo (1995) argue, uncertainty about the expected relative returns of producing for domestic and export markets will affect the composition of output.

Table 1. Variable Descriptions

Variable	Mean ^a	Description
Investment share (<i>I/GDP</i>)	0.007	Share of gross (total) fixed investment in total GDP at 1993 constant prices
GDP tradable goods	0.027	GDP of manufactures and agriculture at 1993 constant prices
GDP nontradable goods	0.018	GDP of construction, electricity, gas and water at 1993 constant prices
Relative price of capital goods	-0.026	Ratio of price index of capital goods over the consumer price index (1993 = 100)
Real interest rate (RIR)	0.004	$((1+i) (1-\pi)^4) - 1$
Domestic nominal interest rate (<i>i</i>)		Money market rate (annualized quarterly). Period average
Inflation (π)		Rate of growth of quarterly CPI (1993 = 100; quarter on quarter)
U.S. real interest rate	-0.002	$[(1+USi) (1-m)] - 1$
U.S. nominal interest rates (USi)		Prime bank lending rate (annualized)
Rate of growth of the import unit value index (<i>m</i>)		Rate of growth of the import unit value index (average quarterly)
Real exchange rate (RER)	-0.001	Real effective exchange rate
Credit to the private sector	0.013	Credit to private sector by deposit money banks and other banking institutions.
Volatility of the real	0.030	Conditional variance of the Garch(1,1) of the <i>International Financial Statistics</i> quarterly real effective exchange rate.
U.S. gross domestic product	0.032	U.S. GDP in billions of chained 1996 dollars
Terms of trade	-0.034	Ratio of Mexico's export and import price indexes (1993=100)
1982 dummy crisis		Dummy variable designed to capture the effect of financial crisis in the Mexican economy. The periods selected correspond to the quarters for which there was a negative rate of growth (<i>y</i> on <i>y</i>). The 1982 crisis covers from 1982:3 to 1983:4, the 1995 crisis from 1995:1 to 1995:4.
1995 dummy crisis		
1982 and 1995 dummy crises		

a. Mean of the rate of growth for all the variables included in regressions, except U.S. and Mexican interest rates (nominal and real), which are in levels.

Note: The dependent variable is the ratio of growth rate of investment to GDP.

Source: National Institute of Statistics, Geography, and Information (INEGI), Mexico; International Monetary Fund, various years, *International Financial Statistics*; Economagic.com [<http://www.economagic.com/>]; HAVER-Emergela; authors' calculations.

If the output multipliers are constant over time, the investment function can be rewritten in terms of growth rates, or in differences of the logs, as follows:

$$(3) \quad g_{I/GDP} = m_T g_T + m_{NT} g_{NT} + g(RIR, g_{p_k}, \sigma_{RER})$$

where *g* denotes the growth rate of the corresponding variable. In this specification the multipliers are analogous to the income accelerators that are standard in empirical investment functions (Rama 1993; Servén 1998).

In a related paper we estimate the investment function described in equation 3 (Lederman and others 2001).⁶ This investment function has as the dependent variable the growth of fixed investment to GDP ratio. The explanatory variables include the growth of tradable and nontradable output, the domestic real interest rate, real exchange rate volatility, variation in the relative price of capital, and the lagged dependent variable.

This basic model was expanded by adding dummy variables for the crises of 1982 and 1995, as well as a term that interacts each crisis with the domestic real interest rate. The aim is to determine whether there is any evidence of “confidence” effects

during crises. A confidence effect would be present in times when increases in interest rates are associated with increases in the growth rate of investment. This phenomenon can theoretically arise when high interest rates signal that the monetary authorities are willing to defend the value of the currency, thus protecting the net worth of firms with liabilities denominated in foreign currency.

Mexican Investment Function and "Confidence" Effects

Results are presented in the first column of table 2. According to the statistic for the *J*-test, the model is well instrumented: the instrumental variables are not correlated with the error term and we can therefore interpret the coefficients as being the impact of the explanatory variables on the dependent variable. In addition, the *Q*-statistics (reported for the first and fourth lags) show no sign of serial correlation and therefore support the model specification.

The results show that a 1 percentage point increase in the growth of tradable output "causes" a 0.96 percentage point increase in the growth of fixed investment to GDP ratio, while an equal increase in the growth of nontradable output would "cause" a 0.30 percentage point increase in the growth of fixed investment to GDP ratio. Thus the multiplier effect of the tradable sector on investment clearly surpasses that of the nontradable sector. Moreover, the *p*-value of the *F*-statistic for the null hypothesis that the two multipliers are equal is 0.01.

The volatility of the real exchange rate has a negative coefficient significant at 5 percent. This coefficient implies that an increase of 1 percent in the volatility growth rate "causes" a reduction of 0.02 percent in the growth of the investment share. Also, the domestic real interest rate has a negative coefficient significant at 5 percent. The estimated coefficient implies that a 1 percentage point increase in the real interest rate leads to a 0.06 percent decline in the growth of investment relative to GDP. The relative price of capital (that is, the ratio of the price index of capital goods over the consumer price index) is not significantly different from zero in this model specification.

The coefficient for the interacted 1982 crisis dummy shows a positive sign but is only significant at 15 percent. The coefficient implies that the overall effect of increases in the real interest rate was positive during this crisis. That is, there is only weak evidence of a confidence effect. The coefficient for the interacted 1995 crisis dummy shows a negative and significant sign. This suggests that there was no confidence effect during this second crisis.

The main conclusion is that we cannot reject the hypothesis of the presence of a structural break in the interest rate coefficient during the 1995 crisis. The coefficients for the 1982 interacted dummy crisis weakly support the existence of a confidence effect.⁷ On the other hand, the 1995 interacted dummy crisis supports the opposite hypothesis, that in this case higher interest rates reduced even more the growth of the investment to GDP ratio.⁸

The model was also extended to explore whether the recovery of Mexican fixed investment growth was due to the financial links with the United States, as argued by Krueger and Tornell (1999), and whether there is an effect of credit availability on

Table 2. Determinants of the Growth Rate of Fixed Investment over GDP in Mexico, 1980–2000

<i>Explanatory variables</i>	<i>Regression 1: Basic model</i>	<i>Regression 2a: With credit rationing</i>	<i>Regression 2b For 1981–94 only</i>
Constant	–0.03 ** (0.01)	–0.02 ** (0.01)	0.07 ** (0.01)
Lagged investment	0.35 ** (0.04)	0.34 ** (0.03)	0.26 ** (0.08)
GDP nontradables	0.30 ** (0.12)	0.13 (0.10)	0.66 ** (0.21)
GDP tradables	0.96 ** (0.14)	1.00 ** (0.17)	0.88 ** (0.27)
Volatility real exchange rate	–0.02 ** (0.00)	–0.03 ** (0.00)	0.00 (0.01)
Relative price of capital	0.04 (0.14)	0.09 (0.13)	–0.43 * (0.25)
Domestic real interest rate	–0.06 * (0.03)	–0.09 ** (0.03)	–0.12 ** (0.05)
Dummy for crises	–0.05 ** (0.02)	–0.08 ** (0.02)	0.04 (0.05)
Crisis 1982*domestic real interest rate	0.13 (0.09)	0.13 * (0.07)	0.46 ** (0.15)
Crisis 1995*domestic real interest rate	–0.29 ** (0.14)	–0.22 ** (0.10)	
U.S. real interest rate		–0.00 (0.03)	
Credit to private sector		0.05 ** (0.01)	
Number of observations	70	70	50
Adjusted R2	0.88	0.84	0.87
R2 standard error of regression	0.04	0.05	0.05
Q-stat at lag 1, p-value	0.64	0.79	0.32
Q-stat at lag 4, p-value	0.51	0.34	0.22
Wald test, p-value	0.01	0.00	0.63
J-test, p-value	0.91	0.90	0.45

** Significant at 5 percent.

* Significant at 10 percent.

Note: The dependent variable is the growth of the ratio of the investment over GDP. Coefficients are from a regression using a generalized method of moments estimator. Variables are the year-on-year differences of their logs, except for the Mexican and U.S. real interest rates, which are in levels. For the first two columns, exogenous variables are U.S. real interest rates, U.S. GDP, terms of trade, four lags of endogenous variables, and seasonal dummy variables. The last column has the same exogenous variables but only two lags of endogenous variables.

Source: National Institute of Statistics, Geography, and Information (INEGI), Mexico.

the rate of growth of the investment share. More specifically, the model explores whether the hypothesis of credit rationing applies to the Mexican experience. Under credit rationing investment could be driven by credit availability for any level of real interest rates (Stiglitz and Weiss 1981). These two questions are explored further in the next section.

Mexican Investment Function and Market Imperfections

To test the hypothesis of credit rationing, we include as an explanatory variable the amount of credit provided to the private sector by financial and nonfinancial

institutions. This variable controls for the quantity of credit but could also reflect variations in supply or demand. To test the Krueger and Tornell effect of financial links between Mexico and the United States we include the U.S. real interest rate.

The results are in the second column of table 2. This model passes the specification tests. The domestic interest rate retains its significant and negative sign even after controlling for credit availability and the U.S. real interest rate. The coefficient for credit to the private sector is also significant and implies that an increase of 1 percent in the growth rate of private credit leads to an increase of 0.05 percent in the growth of the fixed investment share. These results provide evidence of rationing: for a given real interest rate, an increase in credit leads to an increase in the growth of investment to GDP ratio.

As we mentioned above, the domestic real interest rate maintains its significance even after controlling for the U.S. real interest rate (that is, the prime lending rate deflated by the inflation of U.S. import prices). The results show that even though the coefficient for the U.S. real interest rate is negative, it is significant only at 14 percent. Thus there is only weak evidence that access to the U.S. financial market by Mexico's firms operating in the tradable sector was a key feature of the Mexican growth of investment to GDP ratio.

Regarding the other explanatory variables, the multiplier from the tradable sector remains significantly higher than that of the nontradable sector, and the volatility of the real exchange rate presents negative coefficients significantly different from zero. The relative price of capital, on the other hand, has a positive coefficient but is not significant. The coefficient for the interacted term of interest rates with the 1982–83 dummy crisis becomes significant in this specification at 10 percent, thus strengthening the hypothesis that there was a confidence effect during this crisis.

Forecasting the Tequila Crisis and Its Aftermath

The aim of this section is to evaluate the stability of the basic model used to test the existence of a confidence effect by looking at the results for the same regression estimated only for 1981–94. We were unable to test the complete model reported in the second column of table 2 for this sample because of the reduced number of observations. The results are presented in the third column of table 2. This model uses a reduced number of instrumental variables also because of a reduced sample size. Some differences between this and the full sample results: the coefficient for relative price of capital as well as the coefficient for the interacted 1982–83 crisis dummy have the expected sign and are significantly different from zero. Interestingly, the estimated negative coefficient of the real interest rate is larger than the one reported in the first column for results with the full sample. Real exchange rate volatility is not significant and neither is the mean-shifter dummy for the 1982–83 crisis. All the other coefficients retain their expected signs. This specification passes the test for serial correlation and the *J*-test shows that the instrumental variables are not correlated with the error term. The Wald test shows that the coefficient for tradables is not statistically different from that of nontrad-

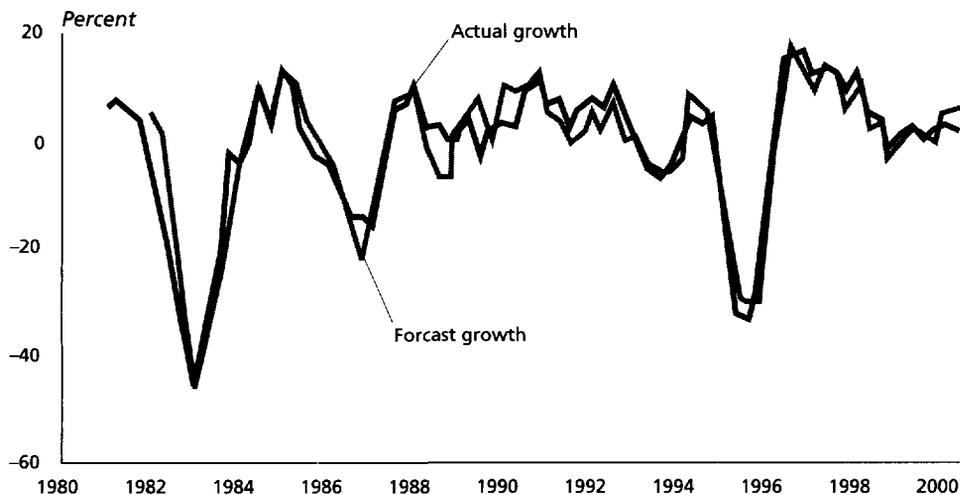
ables for this shorter time period. Hence we conclude that the higher tradables multiplier was present mostly after 1994.

To provide a visual illustration of the capacity of this basic model to explain the behavior of Mexican investment during and after the 1995 crisis, figure 5 presents the forecasted and actual values for the year-on-year growth rate of Mexican fixed investment to GDP ratio. This figure shows that this naïve model quite successfully explains the behavior of Mexican fixed investment. The Theil inequality coefficient comparing the forecast with the actual observations is very low—0.16 (zero indicates a perfect fit). More important, 99 percent of this inequality is due to the covariance between forecast and actual errors (or deviations from the corresponding means). This means that the lion's share of the inequality between the forecast and the actual observations is due to unsystematic error covariance.

Contributions of Explanatory Variables to the Growth of Investment Share

Table 3 shows the contributions of the growth rate of all the significant explanatory variables to the growth rate of investment to GDP ratio during 1995–2000. This exercise is based on the coefficients presented in the second column of table 2 and the growth of the exogenous component of each explanatory variable. The exogenous component of each explanatory variable was derived from the fitted value of each explanatory variable after regressing each one of them on the entire set of instrumental (contemporaneously exogenous) variables used for the generalized method of moments regressions. Adjusted R squares reported in the table show that the fitted vari-

Figure 5. Out-of-Sample Forecast of the Recovery of Mexican Investment Growth, 1980–2000



Source: Authors' calculations; National Institute of Statistics, Geography, and Information (INEGI), Mexico; International Monetary Fund, various years, *International Financial Statistics*.

ables are good predictors of the explanatory variables. This estimation strategy avoids the problem of endogeneity of the explanatory variables in the simulation exercise.

In 1995 the fixed investment to GDP ratio dropped by 28 percent. The contribution of tradables was -1.7 percent while the contribution of nontradables was -2.0 percent. Note that even though the accelerator effect for tradables is bigger than for nontradables, the contribution of tradables to the fall of the growth of investment to GDP ratio was smaller. The reason behind this smaller negative effect of tradables can be explained by the considerable increase of exports during this crisis. An important 6.3 percent of the fall was due to the effect of real exchange rate volatility, while other variables had minor effects.⁹ In 1996 the fixed investment to GDP ratio increased by 9.9 percent. About 6 percent of this growth was due to the reduction in the volatility of the real exchange rate; 7.3 percent to the contribution of tradables; 1 percent to the contribution of nontradables; and 2 percent to credit availability. The relative importance of the contribution of tradables is maintained during 1997–2000. The contribution of real exchange rate volatility was around 1.3 percent (in absolute values), which is not negligible, especially for the observed growth of investment to GDP ratio in the past couple years (3 percent average between 1998–2000), and the contribution of credit availability and domestic interest rates were also important. In contrast, the contribution of variations in the U.S. real interest rates seem to have been quite modest during the past five years.¹⁰

What Was the Dynamic Short-Term Relationship between the Real Exchange Rate and Investment Growth?

Our earlier model does not examine the direct relationship between the real exchange rate and fixed investment growth, because the real exchange rate is

Table 3. Contributions of Selected Variables to the Growth of the Ratio of Fixed Investment to GDP in Mexico, 1995–2000

(percent)

Variable	1995	1996	1997	1998	1999	2000
Growth of share of fixed investment in GDP	-28.0	9.9	12.6	5.4	2.0	3.0
Contribution of nontradables	-2.0	1.0	0.9	0.3	0.5	0.9
Adjusted $R^2 = 0.80$	(-15.3)	(7.6)	(6.9)	(2.4)	(3.6)	(6.9)
Contribution of tradables	-1.7	7.3	7.1	5.3	4.5	6.8
Adjusted $R^2 = 0.75$	(-1.7)	(7.3)	(7.1)	(5.3)	(4.5)	(6.8)
Contribution real exchange rate volatility	-6.3	6.3	1.5	-1.6	1.4	-0.5
Adjusted $R^2 = 0.80$	(211.4)	(-210.9)	(-51.1)	(54.3)	(-47.1)	(16.2)
Contribution of domestic real interest rate	-0.6	0.1	-0.2	-1.0	-1.4	-0.5
Adjusted $R^2 = 0.76$	(1.8)	(-0.6)	(2.4)	(10.8)	(15.4)	(5.0)
Contribution of U.S. real interest rate (exogenous variable)	0.0	0.0	0.1	0.1	0.0	-0.0
	(6.3)	(6.5)	(13.2)	(15.6)	(2.5)	(-4.9)
Contribution of credit	-0.7	-2.0	0.6	0.5	-0.4	-0.8
Adjusted $R^2 = 0.93$	(-14.0)	(-40.0)	(1.2)	(10.5)	(-7.9)	(-16.2)

Note: The numbers in parentheses are the growth rates of the exogenous component. The contributions are calculated as the estimated coefficient (from the second column of table 2) times the growth rate of the exogenous component of each variable. The adjusted R^2 corresponds to the regression of each explanatory variable on the set of instruments used in the generalized method of moments estimation.

Source: National Institute of Statistics, Geography, and Information (INEGI), Mexico.

expected to affect investment only indirectly, through its effects on the relative price of capital, the composition of output, and economic volatility. Here we use the impulse response function derived from vector autoregressions to study the relationship between fixed investment growth and the real exchange rate.

We ran a vector autoregression with the growth rates of fixed investment, the relative price of capital, and the real exchange rate as endogenous variables. The choice of these variables was inspired by evidence for Mexico that the relative price of capital may be an important channel through which variation in the real exchange rate (and its determinants, such as terms of trade) affect domestic investment (Warner 1994), although our previous results do not support this conclusion.¹¹ We also included the following exogenous variables likely to affect the three endogenous variables: the growth of U.S. GDP, variation in public consumption, variation in the terms of trade, the U.S. real prime lending rate, dummy variables for openness and NAFTA, and a constant.

The corresponding impulse response function assumes that the real exchange rate is more exogenous than the relative price of capital, which is assumed to be more exogenous than fixed investment. Twelve lags of the endogenous variables were included in the vector autoregression, because the Akaike, Schwartz, and log likelihood tests all indicated that this was the best distributed lag specification when compared with one, four, and eight lags. The impulse response function illustrates the effect on investment growth of a one standard deviation innovation in the real exchange rate. The results show that the response of investment growth to an appreciation of the real exchange rate is positive at first but becomes negative and significant during the second year, after about five quarters (figure 6).¹²

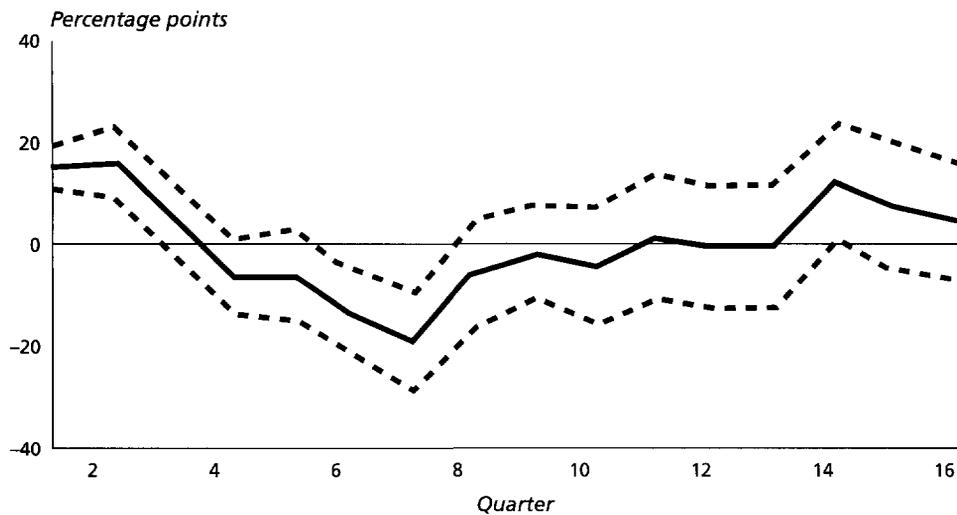
The finding that an appreciation of the real exchange rate in Mexico is associated with economic expansion in the short run is consistent with the conclusions of other recent empirical studies. Kamin and Rogers (1998), for example, also find that appreciation has a positive effect on output—and depreciation a negative effect. But they argue that there are multiple channels through which this effect takes place, including its impact on government spending, monetary aggregates, and capital flows. In fact, Kamin and Rogers only tentatively reject the hypothesis that the channel is the inflation rate.

Four quarters after the simulated appreciation, the growth of fixed investment becomes negative. A plausible interpretation of these dynamic effects is that appreciation has an initially positive effect that could be due to income effects (or positive net worth effects) and the reduction of the relative price of capital. Later, as a substitution effect takes hold (see below), investment growth declines as a result of the real appreciation of the currency.

Did Real Exchange Rate Variation Have Short-Term Substitution Effects?

Based on pairwise vector autoregressions, the impulse response functions presented in figure 7 look at the effect of the real exchange rate on the relative growth of the

Figure 6. Response of Fixed Investment to a One Standard Deviation Innovation in the Real Exchange Rate



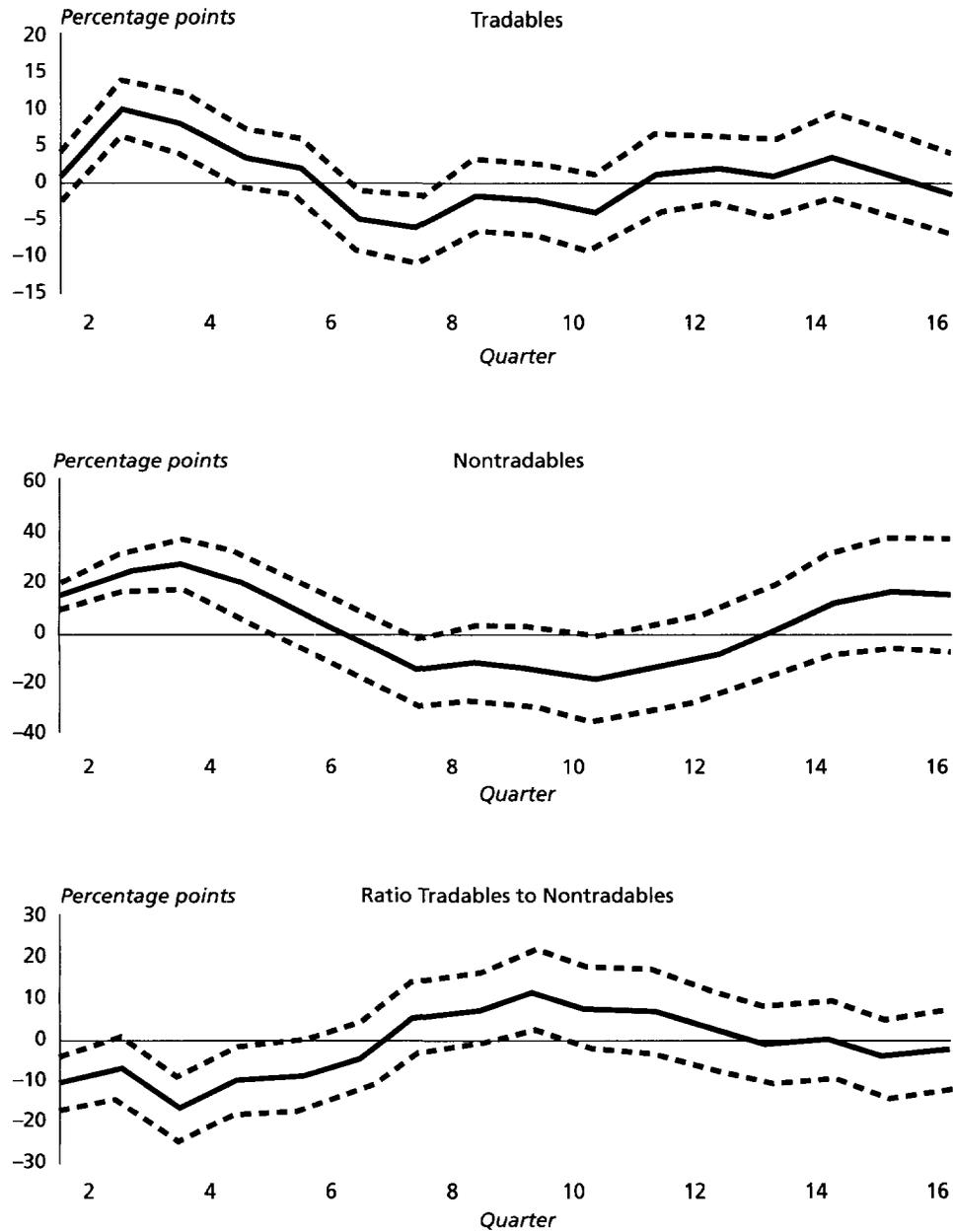
Note: Variables are in the year-on-year differences of the logs. The vector autoregression includes as endogenous variables the real exchange rate, relative price of capital, and fixed investment. The exogenous variables are U.S. GDP, the U.S. real interest rate, public consumption, dummy variables for openness and Nafta, terms of trade, and a constant.

Source: Authors' calculations; National Institute of Statistics, Geography, and Information (INEGI), Mexico; International Monetary Fund, various years, *International Financial Statistics*.

tradable and nontradable sectors. The model was estimated with 1, 4, 8, and 12 lags, but the figure shows the impulse response functions based on the lag specification supported by the Akaike criterion. The endogenous variables are the real exchange rate and the growth rate of tradables, of nontradables, or of the ratio of tradables to nontradables; the real exchange rate is the more exogenous. The exogenous variables are U.S. GDP, the U.S. real interest rate, public consumption, terms of trade (all the variables are the first difference of their logs), and dummy variables for openness and NAFTA.

The first two impulse response functions show that real exchange rate appreciation tends to raise the growth rate in the tradable and nontradable sectors in the short run, an effect lasting a few quarters. This effect is larger for the nontradable sector.¹³ The third impulse response function provides evidence that real exchange rate variation has significant substitution effects: an appreciation slows the growth in the ratio of tradables to nontradables. This effect is also clear from figure 8, which shows the growth in the ratio of tradable to nontradable output accelerating after the depreciations during the debt and tequila crises. In other words, while real exchange rate depreciation tends to have a short-run negative income effect associated with a decline in the growth rates of tradable and nontradable output, this effect is larger for nontradables. We interpret this difference in the size of the effect as evidence of a temporary substitution effect. And this effect implies a permanent change in the composition of Mexican output in favor of tradables.

Figure 7. Response of Growth in Tradables, Nontradables, and the Ratio of Tradables to Nontradables in Mexico to a One Standard Deviation Innovation in the Real Exchange Rate



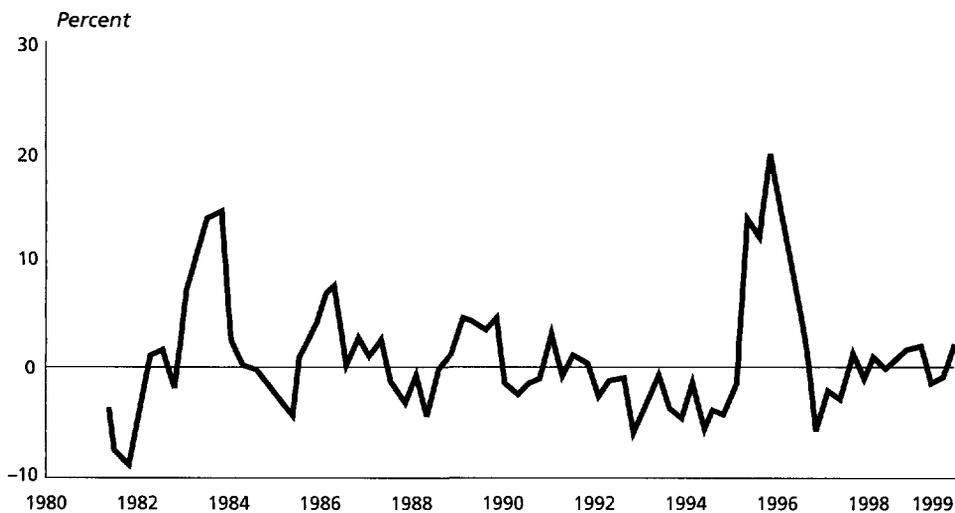
Note: Variables are in the year-on-year differences of the logs. The exogenous variables are U.S. GDP, the U.S. real interest rate, public consumption, dummy variables for openness and Nafta, terms of trade, and a constant. *Source:* Authors' calculations; National Institute of Statistics, Geography, and Information (INEGI), Mexico; International Monetary Fund, various years, *International Financial Statistics*.

Conclusion

Since the eruption of the Mexican tequila crisis of 1995 and the Asian crises of 1997, much has been written about what causes financial crises in developing countries. Much less attention has been given to what happens in the productive economy after these crises.¹⁴ This analysis has identified the main macroeconomic factors behind the recovery of the Mexican economy following the 1995 crisis:

- An increase in export growth and a decline in import growth during the crisis played a crucial role in cushioning the fall in GDP. During the recovery import growth rose, slowing GDP growth, and export growth, though still important, contributed less to GDP growth than during the crisis.
- Export revenues (in constant pesos) grew more rapidly during and after the crisis. The rise in growth was not all due to the “accounting” effect of the real exchange rate. It was also due to the acceleration in the growth of the volume of non-oil exports in this period. The competitiveness of the tradable sector improved: as the real exchange rate depreciated, the manufacturing remuneration rate in U.S. dollars fell, remaining at that low level during 1995–99. Since 1994 unit labor costs have also tended to fall, even when the real exchange rate appreciated.
- Fixed investment growth plummeted during the crisis but recovered sharply in 1996, reaching rates higher than those during the precrisis period. The contribution of fixed investment was thus critical in the postcrisis period.

Figure 8. Average Quarterly Growth in the Ratio of Tradable to Nontradable Output in Mexico, 1980–99



Source: Authors' calculations; National Institute of Statistics, Geography, and Information (INEGI), Mexico; International Monetary Fund, various years, *International Financial Statistics*.

- A basic investment model, differing from the standard model only by the assumption that tradable and nontradable output growth have different multiplier effects on investment, does quite well in predicting the recovery of investment in 1996–2000.
- The model for the whole period finds that the tradable sector has a larger multiplier effect than the nontradable sector. Real exchange rate uncertainty was also a significant determinant of fixed investment.
- Both the amount of credit and the real interest rate seem to be important determinants of the growth of fixed investment to GDP ratio. These results provide evidence of rationing: for a given real interest rate, an increase in credit leads to an increase in the growth of investment to GDP ratio. The impact of real exchange rate variation on fixed investment was initially dominated by income effects, but substitution effects were also present, with real exchange rate depreciation producing a larger decline in the growth of nontradable output than in that of tradable output. Thus the real exchange rate depreciation in 1995 had a negative income effect associated with a decline in the growth of tradable and nontradable output in the short run. But because of the substitution effect the decline was larger for nontradables, and the growth of tradables relative to nontradables accelerated after the depreciation.

From these findings emerge several important empirical conclusions: The decline and subsequent rise in GDP growth were primarily due to the behavior of fixed investment. Fixed investment growth was negative during the crisis because the short-lived hike in interest rates did not positively affect investment—there was no confidence effect. Also, investment declined because of the negative effect of higher volatility in the real exchange rate. The real depreciation also had a negative income effect that influenced investment decisions. The recovery of fixed investment observed during the year after the crisis was linked to the substitution effect of the real depreciation in favor of the tradable sector. In turn, the growth in the tradable sector, with its relatively large multiplier effect, drove the recovery of investment. Indeed, the real depreciation produced a notable improvement in the competitiveness of Mexican manufactures, reflected in the increase in the growth of the volume of non-oil exports. This may explain why growth in tradables declined less than growth in nontradables in 1995.

The findings also point to a policy implication: During currency crises there are tradeoffs related to interest rate defenses of the currency. High nominal interest rates used to contain inflation are associated with higher real interest rates, which then contribute to a downturn in investment. But depreciation of the currency can have adverse consequences for investment in the short run as a result of its effects on economic uncertainty and its negative income effect. In the medium term, however, a real depreciation is healthy because it stimulates growth in exports and increases the share of output from the tradable sector, which then has a large multiplier effect on investment. At the very least, we have shown that the Mexican currency crisis of 1995 is not a case in which high real interest rates benefited the economy.

Notes

1. For example, an IMF report (1996) states that “in Mexico, tight fiscal and monetary policies helped reduce actual and expected inflation in the first half of [1995], contributing to further gains in confidence, declines in interest rates, and the stabilization of the peso” (p. 25).

2. Krueger and Tornell (1999) suggest that the increase in export growth was facilitated by previous structural economic reforms, such as privatization, deregulation, removal of trade barriers, and the entry into the North American Free Trade Agreement (NAFTA).

3. Any increase in trade barriers to diminish negative short-term growth effects during the crisis would have negatively affected Mexico’s partnership in NAFTA.

4. The rate of change in export revenues measured at constant local prices can be decomposed into a relative price or “accounting” effect and a volume effect: $\dot{x} = R\dot{E}R + \dot{Q}_X$. The accounting effect refers to the fact that a real depreciation raises the value of exports in domestic currency relative to the general consumer price level.

5. See section II in Lederman and others (2001) for a model of optimal investment rules that gives theoretical grounds for this assumption. In brief, $m_T > m_{NT}$ because of the higher capital intensity of production in the tradable sector.

6. We refer to the main findings of Lederman and others (2001) where appropriate. The empirical analyses in that paper rely on publicly available data, primarily from the Mexican statistical agency (INEGI) and the International Monetary Fund. The data have a quarterly frequency and cover the period from 1980 through the second quarter of 2000. The econometric models were estimated using the generalized method of moments estimator, which controls for the joint endogeneity of the explanatory variables.

7. However, we acknowledge that this break may be due to other effects that are not really “confidence” effects. For example, the banking system was nationalized during that time, and the change in the coefficient’s sign during that crisis may be due to interest-rate controls.

8. We also acknowledge that we cannot test for “confidence” effects with a duration of less than one quarter, given the frequency of the available data.

9. The sum of the contributions from all the variables reported does not equal the growth of fixed investment because there are other (omitted) determinants of investment.

10. The table does not include contributions of relative price of capital because this variable was statistically zero in this model.

11. The real exchange rate used here is the International Monetary Fund’s real effective exchange rate.

12. Among the exogenous variables, the growth of U.S. GDP has a positive and significant coefficient (6.2; t -statistic of 2.4). The U.S. real interest rate is also significant (−0.4; t -statistic of −2.6).

13. Among the exogenous variables, only the terms of trade has a significant coefficient in the regression for the growth of tradable output. The U.S. GDP growth rate has a positive coefficient (0.8), but its t -statistic is low (1.4). None of the exogenous variables was a significant determinant of nontradable output growth.

14. See Perry and Lederman (1999) for a comparison of the aftermath of the Latin American and Asian crises in the 1990s.

References

- IMF (International Monetary Fund). Various years. *International Financial Statistics*. Washington, D.C.
- . 1996. *World Economic Outlook*. Washington, D.C.

- Kamin, Steven, and John Rogers. 1998. "Output and the Real Exchange Rate in Developing Countries: An Application to Mexico." Board of Governors of the Federal Reserve System International Finance Discussion Paper 580. Washington, D.C.
- Krueger, Anne, and Aaron Tornell. 1999. "The Role of Bank Restructuring in Recovering from Crisis: Mexico 1995–98." NBER Working Paper 7042. National Bureau of Economic Research, Cambridge, Mass.
- Lederman, Daniel, Ana María Menéndez, Guillermo Perry, and Joseph E. Stiglitz. 2001. "Mexican Investment after the Tequila Crisis: Basic Economics, 'Confidence' Effects or Market Imperfections?" World Bank, Latin America and the Caribbean Region, Chief Economist's Office, Washington, D.C.
- Maloney, William, and Rodrigo Acevedo. 1995. "Trade Reform, Uncertainty and Export Promotion: Mexico 1982–88." *Journal of Development Economics* 48 (October): 67–89.
- Perry, Guillermo, and Daniel Lederman. 1999. *Adjustments after Speculative Attacks in Latin America and Asia: A Tale of Two Regions?* Washington, D.C.: World Bank.
- Pindyck, Robert. 1988. "Irreversible Investment, Capacity Choice and the Value of the Firm." *American Economic Review* 78 (December): 969–85.
- Rama, Martín. 1993. "Empirical Investment Equations for Developing Countries." In Luis Servén and Andres Solimano, eds., *Striving for Growth after Adjustment: The Role of Capital Formation*. Washington, D.C.: World Bank.
- Servén, Luis. 1998. "Macroeconomic Uncertainty and Private Investment in Developing Countries: An Empirical Investigation." Policy Research Working Paper 2035. World Bank, Development Research Group, Washington, D.C.
- Stiglitz, Joseph E., and Andrew Weiss. 1981. "Credit Rationing in Markets with Imperfect Information." *American Economic Review* 71 (June): 393–410.
- Warner, Andrew M. 1994. "Mexico's Investment Collapse: Debt or Oil?" *Journal of International Money and Finance* 13 (April): 239–56.

Comment on “The East Asian Crisis—Two Years Later,”
by Eisuke Sakakibara, and “Mexico—Five Years after
the Crisis,” by Daniel Lederman, Ana María
Menéndez, Guillermo Perry, and Joseph E. Stiglitz

Alejandro Werner

From the experiences of recent crises in emerging markets, Eisuke Sakakibara (East Asia) derives tentative conclusions about reform of the international financial architecture and Daniel Lederman and his coauthors (Mexico) draw tentative conclusions about the sources of recovery and about crisis management. The authors take very different approaches: Sakakibara covers a broad range of topics in a speculative manner, while Lederman and his coauthors perform a broad quantitative study of the sources of growth in Mexico following the 1995 crisis and a thorough econometric analysis of the investment function.

I turn first to the issues Sakakibara raises about the nature of recent crises and the options emerging economies have in defending themselves against the instability of international capital markets.

Solutions to Crisis through Domestic Reform and the International Financial Architecture

In the time that has elapsed since the recent crises, a consensus has been emerging on their causes. Essentially, in every crisis country we have seen a mix of weak fundamentals (macro- and microeconomic) and liquidity issues that opened the door to self-fulfilling panics. The relative importance of these two elements depends on the country. But even for the hypothetical case in which the crisis is a pure financial panic, four points should be highlighted:

- Sakakibara claims that macroeconomic fundamentals in most of the affected countries were strong before the Asian crisis. But even if macroeconomic imbalances are not the cause of a crisis, once the economy suffers a financial collapse the public sector will probably end up paying a large share of the bill. Fiscal retrenchment and tight monetary policy are therefore in order.
- As Sakakibara notes, rapid financial liberalization is one of the main causes of the large buildup of short-term liabilities to nonresidents. But we should

Alejandro Werner is director of economic studies at the Bank of Mexico.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

not overlook the fact that poor corporate governance and weak financial systems also bias capital inflows toward the short term. Thus corporate governance reform also helps on the macroeconomic front, by lengthening the maturity of capital inflows and increasing the share of non-debt-creating flows.

- The Mexican experience with a floating exchange rate shows that the private sector—financial and nonfinancial—is taking exchange rate risk into account, and today we no longer see the currency mismatches that we saw in the past. For example, in 2000 about 80 percent of total corporate foreign currency debt was held by highly exporting firms whose annual exports to total foreign currency debt ratio was 70 percent and who have large holdings of foreign currency assets. This supports the claim made by several economists (see Obstfeld 1998 and Burnside and others 1999) that a fixed exchange rate regime works like an implicit guarantee and thus contributes to the buildup of a large short position in foreign currency.
- Although I agree with Sakakibara that market participants do not behave rationally, his arguments for multiple equilibria do not require irrationality: all the models with multiple equilibria assume rational expectations.

For emerging markets facing the problem of self-fulfilling panics, the only solution that will allow these countries to benefit from short-term inflows, and permit them to perform a maturity transformation function similar to that performed by banks, is to have an international lender of last resort. Improving the International Monetary Fund's contingent credit line facility—which, as it stands, is proving to be inoperative—could go part way toward establishing an international lender of last resort.

Of the two other solutions proposed by Sakakibara, controls on capital flows and a do-it-yourself lender of last resort, I would argue for the second, for several reasons. First, short-term capital flows perform a very useful role as a disciplining device. Second, permitting such flows has been the first step in attracting capital that might eventually become medium-term investment. And finally, there is no evidence that capital controls work.

I completely agree with Sakakibara on the need to deepen financial markets and strengthen regional trade and financial links. But we need to remember that corporate governance must also be improved to develop long-term capital markets.

Sakakibara also proposes creating regional defensive mechanisms—regional monetary funds and exchange rate regimes. In theory, the optimal solution to the problem of a financial panic is to have a lender of last resort. The decision on whether it should be regional or global depends more on politics than on economics. But we should bear in mind that contagion has been a regional phenomenon, and that having a regional lender of last resort could exacerbate it. Several countries would be likely to approach a regional lender of last resort for funds at the same time, weakening its position.

The Brazilian, Mexican, and Asian experiments with floating exchange rate regimes are working quite well. Thus short of having well-established regional cur-

rency unions, we should concentrate on deepening financial systems and debt markets and building sufficient liquidity reserves to support these regimes.

Sources of Recovery and Crisis Management

In investigating the factors behind the recovery of the Mexican economy after the 1995 crisis, Lederman and his coauthors first document the importance of investment in explaining both the recession and the recovery. This line of analysis justifies the aim of their article, which is to understand the main determinants of investment. But the data the authors present hold other interesting patterns that are not discussed.

In particular, the data show that consumption is more important than investment in explaining the recession, but its contribution to the upturn is relatively small (see figure 2 in the article). This observation points toward significant asymmetries in the factors explaining the pattern of recession and recovery. These asymmetries might be explained by the credit crunch in Mexico after 1995, which affected different sectors of the economy in different ways.

The authors also emphasize the behavior of exports: in 1996–99 non-oil exports from Mexico grew at an average annual rate of more than 15 percent. But the authors' analysis of the main determinants of Mexican exports omits the impressive economic growth that the United States has experienced, with an average growth rate of 4 percent over the past four years.

Lederman and his coauthors do an interesting job in estimating an investment function for Mexico, focusing on the different multiplier effects of the tradable and nontradable sectors and the role of the relative price of capital. Another interesting contribution is the evidence they find for the segmentation that has taken place in Mexico between firms that can fund themselves in international markets and those that lack access to credit because of the crisis in the Mexican banking system. Thus it is surprising that the authors are unable to find evidence in support of the existence of credit constraints, though other studies have found results supporting this hypothesis (for example, Copelman 2000). In their implementation of the investment function it would have been interesting to allow for the impact of U.S. economic growth and the increase in U.S. foreign direct investment in Mexico (total U.S. investment abroad has doubled since 1994).

Another important omission that might have biased the authors' results is their failure to control for country risk and access to U.S. markets. Since these two variables are highly correlated with the real exchange rate, this omission might have led the authors to place excessive importance on the role played by the exchange rate during the recovery. In 1996–98 investment in Argentina followed a pattern similar to that observed in Mexico, but without real exchange rate depreciation, suggesting that the role of the real exchange rate might have been smaller than Lederman and his coauthors show. Supporting this conclusion are the results from the authors' impulse response functions relating to the impact of a real exchange rate shock on investment and on tradable and nontradable output. This impact looks a lot like the effect of a positive capital inflow shock. Thus the authors' results could have been

driven by the omission of capital flows rather than by an exogenous depreciation of the real exchange rate.

Finally, the authors might have considered using private rather than total investment in their analysis, given the significant changes in public investment in Mexico during the period covered. Public investment fell from an average of 8 percent of GDP in 1980–82 to 2.6 percent in 1997–99.

Interest Rate Defenses and Crisis Management

Lederman and his coauthors find that high real interest rates lead to lower investment. When they introduce a dummy variable for the crises the Mexican economy suffered in 1982 and 1995 and interact it with the interest rate, they find evidence of credibility effects for 1982 but not for 1995. Thus they conclude that in the aftermath of the 1995 crisis interest rate hikes were not helpful in reestablishing confidence and promoting recovery.

But the statistical evidence they present does not support this conclusion. Nor is it sufficient for criticizing interest rate defenses of exchange rates, for several reasons:

- The argument made in the literature (see Goldfajn and Gupta 1999) in support of interest rate hikes is that once the nominal exchange rate moves beyond a certain level, there will be a large, rapid pass-through to inflation and a deterioration of corporate and bank balance sheets. These phenomena will erase any other real exchange rate benefit. This will force the authorities to substantially raise real interest rates for a long period, in an effort to avoid an excessive depreciation and, therefore, to avoid bankruptcies and an inflationary outburst. Thus failing to find evidence that the dampening effect of interest rates on investment declines during the 1995 crisis does not prove that the Mexican authorities were not able to avoid a bigger burst of inflation and an even larger future increase in interest rates. More appropriate for proving this hypothesis is the empirical work of Goldfajn and Gupta (1999), who do find partial evidence in favor of interest rate hikes as a mechanism for limiting exchange rate depreciations or for inducing fast exchange rate reversals. A comparison of the experiences of the Asian countries and Brazil with that of Mexico points toward the benefits of containing the inflationary effects of an uncontrolled devaluation.
- There is an important measurement problem in the construction of real interest rates for Mexico, where the bursts of inflation led to important differences between ex post and ex ante interest rates due to the large differences between expected and observed inflation. Lederman and his coauthors construct real interest rates using observed past inflation. As a result, and because of the increases seen in the Mexican inflation rate in 1982 and 1995, real interest rates were negative during the crisis periods. What we see in the regression is that in 1982, because of financial repression, real interest rates were extremely negative (see figure 11 in the article) but investment fell significantly (rather than increasing). So the impact of interest rates was smaller than usual, but not due to an increase in confidence.

- Not all interest rate increases represent monetary policy shocks, as the authors assume. Thus it should be interesting to assess the impact of high interest rates on investment when controlling for some measure of country risk. That would make it possible to see the difference between simply matching an increase in country risk and increasing the spread between domestic and foreign interest rates.
- It is very important to adjust for public investment, which fell by 50 percent in 1982–83.

To conclude, the study of the determinants of investment, the role of the exchange rate, and the role of foreign interest rates is extremely important for understanding the real effects of the crisis and the process of recovery. However, the authors should have left the policy debate on interest rate defenses for future work and stayed with the main objective of their empirical work: simply to assess whether the recovery of Mexican investment after the tequila crisis can be explained with real variables.

References

- Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo. 1999. "Hedging and Financial Fragility in Fixed Exchange Rate Regimes." NBER Working Paper 7143. National Bureau of Economic Research, Cambridge, Mass.
- Copelman, Martina. 2000. "Financial Structure and Economic Activity in Mexico." Instituto Tecnológico Autónomo de México, México, D. F.
- Goldfajn, Ilan, and Poonam Gupta. 1999. "Does Monetary Policy Stabilize the Exchange Rate Following a Currency Crisis?" IMF Working Paper 99/42. International Monetary Fund, Washington, D.C.
- Obstfeld, Maurice. 1998. "The Global Capital Market: Benefactor or Menace?" *Journal of Economic Perspectives* (fall): 9–30.

*Corporate
Governance and
Restructuring—
Lessons from
Transition and
Crises*

Ownership Structure, Legal Protections, and Corporate Governance

I. J. Alexander Dyck

This survey presents a functional framework for understanding corporate governance issues and examines the theoretical and empirical evidence on the effect of legal protections (legal rules protecting investor rights) and the ownership structure of firms on governance. Legal protections can improve governance by shifting power to investors and lowering the cost of resolving disputes. The effectiveness of legal protections, though, depends on complementary institutions that provide enforcement and ensure the provision of information for investors. Where such complementary institutions are inadequate, such alternative approaches as ownership structure may provide a lower cost approach to provide the functions of governance. Evidence for established and privatized firms suggests that there are links among legal protections, ownership structures, and firms' financial performance. The article draws implications about policies on ownership structure in privatized and established firms and identifies issues to be considered in reforming governance systems.

There are costs to using the market to organize transactions; entrepreneurs create firms to alleviate these costs. In so doing, the authority of the entrepreneur substitutes for the price mechanism in the allocation of resources. This paper focuses on corporate governance institutions that influence the allocation and exercise of entrepreneurial authority.

In this article I don't approach governance from an institutional or neoclassical economic perspective. An institutional approach to governance takes existing institutions, such as the bank-centered corporate governance systems of Germany and

I. J. Alexander Dyck is associate professor at the Harvard Business School. Parts of this article draw on Dyck (2001). For helpful discussions, the author thanks Erik Berglöf, Harry Broadman, David Ellerman, Avner Greif, Roumeen Islam, Carl Kester, Rafael La Porta, Jay Lorsch, Thomas McCraw, David Moss, John Nellis, Katharina Pistor, Julio Rotemberg, Bruce Scott, Mary Shirley, Joseph E. Stiglitz, Philip Wellons, participants in the Global Corporate Governance Initiative of the Harvard Business School, and seminar participants at the Harvard Business School and the World Bank. The author also gratefully acknowledges the support of the World Bank and the Harvard Business School's Division of Research.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

Japan, as given and asks how these institutions could produce the services they offer more effectively. A neoclassical economic approach avoids discussing institutions and concentrates almost exclusively on assumed functions, exploring the limitations imposed by fundamental variables such as technology, endowments, and preferences.¹ In the relatively rare circumstances in which institutions are introduced, they are viewed solely as imposing constraints on sophisticated investors.

The approach I use here takes the functions of a governance system as given and asks how different institutional arrangements can address these functions.² The approach thus shares the economic focus on functions, but it does not assume that institutional differences either do not affect outcomes or serve solely as constraints. Instead, it admits the possibility that institutions can complement markets by lowering the costs of providing core functions. This functional approach does not presume that there is only one institutional way to address governance concerns. It forces an examination of the range of institutions and policies involved in governance—a range much greater than that suggested by the institutional perspective. The approach is well suited to examining the possibilities for reform.

A Framework for Understanding Corporate Governance

The transaction at the heart of corporate governance is an investment in a corporation that is simultaneously met not with payment but with a promise of future returns. Correspondingly, I define a corporate governance system as the complex set of socially defined constraints that affect expectations for how authority in firms will be exercised—and thus how the system affects the willingness to make investments in corporations in exchange for promises.³ These constraints include those taken as exogenous by investors in firms when they make their investments (what I call institutions of governance, such as the legal system) and those introduced by investors (what I call policies, such as the compensation system for managers).

This view of governance encompasses many investments in exchange for promises based on explicit contracts and implicit contracts. Investments of equity and debt capital are obviously corporate governance transactions—payment is by definition delayed. Workers and management also make investments with delayed payment. When workers are asked to invest in skills of use only to a particular firm (rather than in general skill formation) and payment for that investment is deferred, they are engaged in corporate governance transactions. Managers also make investments in exchange for promises, as they often invest in specific skills and knowledge and put their reputations at risk with only limited security that they can retain their positions and have the discretion to take actions that reveal their ability.⁴

I use the term *promise* to emphasize both the separation between the quid and the quo⁵ in governance transactions and the fact that such exchanges are often based on implicit understandings in addition to explicit contracts. With any investment transaction there are natural limits to the ability to completely specify actions under all future contingencies. In addition, the terms of the initial contract may themselves be violated. When making an investment, the investor therefore calculates expected

returns on the basis of not only specific terms in a contract but also implicit understandings and the expected outcome from inevitable competing claims on the wealth generated by the firm.

Functions Provided by a Corporate Governance System

Identifying the primary function of corporate governance as facilitating investment helps clarify thinking; it is less useful when we attempt to link institutions to the provision of functions. More helpful is the identification of core functions that contribute to this primary objective. Merton and Bodie (1995) identify the core functions performed by the financial system, a system that shares many features of a corporate governance system: clearing and settling payments, pooling resources and subdividing shares, transferring resources across time and space, managing risk, providing information, and dealing with incentive problems. With the exception of clearing and settling payments, these functions are also core functions of corporate governance systems.

Since many investments are beyond the capacity of individual investors, corporate governance institutions provide a way to pool resources. Investments are also more likely when risks associated with those investments are pooled, shared, and allocated to those with greater risk-bearing capacity. Corporate governance institutions allocate risks. Providing information and managing incentive problems are correctly placed at the heart of analyses of governance systems, because asymmetries in information discourage investments by increasing the probability of adverse selection and moral hazard. Corporate governance institutions determine the extent and timeliness of information disclosure as well as the incentives to engage in monitoring.

But perhaps the defining function of corporate governance institutions relates to the authority held by the entrepreneur. Corporate governance systems determine how competing claims on the wealth generated by the corporation are to be resolved. The entrepreneur in the firm is not the only agent with authority. In both the approach of Williamson (1985) and the more formal approach summarized by Hart (1995), the incomplete contract literature has focused attention on how ownership of physical assets brings authority. When the contract is silent, the owner of the assets has the residual right to specify how these assets are distributed. Also, the government has authority to allocate power to the entrepreneur or to such stakeholders as labor in disputes over allocation of wealth, such as bankruptcy proceedings. In short, the costs of resolving competing claims on the wealth generated by the corporation depend on the identity of those who have power in the firm and in society more generally.

Good Governance Defined

In all economies in which investments are exchanged for promises, mechanisms are in place to provide these core functions. A “good” governance system is distinguished from a weak governance system by its capacity to provide these functions at a lower cost. A good governance system supports a continual process of mobilizing

scarce resources, including investments by financiers, trade credit by suppliers, and specific human capital by labor and management, and allocating them to their most promising uses. Thus at the most aggregate level, the measure of the strength of a governance system is the extent and efficiency of investments in corporations.

An alternative measure of the strength of a governance system is the efficiency with which the core functions are handled. A good governance system allocates risk to those with a comparative advantage in bearing risk. It reduces information asymmetries between insiders and outsiders and among outside providers of resources. Indicators of effectiveness include the availability of information, adherence to accounting standards, and the precision of estimates of returns. A good governance system addresses inevitable incentive problems. Insiders are rewarded for performance that realizes outsiders' objectives. More critically, insiders are removed rapidly when evidence emerges of weak management or inappropriate practices. A good governance system minimizes the ability to divert returns and the costs of resolving inevitable disputes over the wealth generated by the corporation. Measures of the ability to divert returns include the extent to which a controlling shareholder is willing to pay a premium for control that cannot be accounted for by the shareholders rights to firm cash flow. Measures of costs of dispute include the incidence of disputes and the speed and reliability of dispute resolution.

Institutions of Corporate Governance Systems

Both governance institutions and governance policies (discussed below) affect beliefs about whether promises will be fulfilled and influence expected future returns—and in so doing influence investment choices. Institutions are determined at the societal level and are taken as exogenous by investors in making their decisions.⁶ Policies are under the control of at least some investors.

NATIONAL LEGAL PROTECTIONS. A claim associated most closely with the pioneering work of La Porta and others (1998) is that national legal protections have a profound influence on the willingness to invest resources in firms and are a central institution of corporate governance. This approach has emerged both from the theoretical work that emphasizes the limits to contracting and from the observation that no country relies solely on the simple legal solutions offered by contracting. Everywhere governments introduce additional legal mechanisms that specify procedures and remedies for the most common breakdowns in governance.

A key assumption of this approach is that the default contract provided by national law can improve outcomes by affecting the bargaining over competing claims on the wealth generated by the corporation—and consequently the willingness to invest in firms. Echoing a theme found in the new institutional economics, the institution of legal rules is a constraint that can complement markets rather than limiting market transactions.

What legal rules are important for lowering the costs of resolving competing claims? First, the research of La Porta and others (1998) focuses on the elements of

corporate law that favor minority shareholders, rather than management or a dominant shareholder, in corporate decisionmaking. The vast majority of these mechanisms provide for little interference with management during normal times but significant interference with low transactions costs when a situation deteriorates. Specifically, La Porta and others focus on six anti-director rights to indicate who holds power in firms: the right to vote by proxy through the mail, a ban on blocking sales of shares before meetings, the use of cumulative voting, the existence of oppressed minority rights, the existence of a preemptive right to new issues for existing shareholders, and a low threshold for calling an extraordinary meeting. They also collect information on the presence or absence of two additional provisions in the company code: a mandatory dividend and one-share-one-vote rules that link cash flow rights to voting rights, both of which could protect the interests of minority shareholders.

A shared feature of these legal protections is their clear allocation of power. Legal protections temporarily concentrate control and provide a credible threat to replace insiders, be they managers or controlling shareholders. In addition to shareholder meetings, extraordinary actions hinge on such protections as class action lawsuits and takeovers. In a class action lawsuit a group of equity investors seeks to sidestep the board and directly stop current management actions or actions approved by the board. In a takeover current controlling investors are replaced with a new controlling investor who can address problems in the current ownership structure and take more forceful action in dealing with management. Takeovers help overcome the public good problem associated with monitoring management that arises when there are a large number of small shareholders. But these efforts to temporarily increase ownership concentration can be very expensive and are therefore best viewed as extraordinary measures.

Second, La Porta and others focus on the interests of providers of debt finance, which is far more important than equity finance in most developing economies. They concentrate on rules that clearly allocate power in the case of competing claims on the wealth created by the firm. Bankruptcy laws specify criteria for determining when promises have not been kept and identify procedures for reallocating control over the use of assets and the distribution of assets (normally giving control temporarily to a judge, who often transfers it to a trustee controlled by creditors). When these protections are strong, the costs of invoking them are low and the bargaining process to realize a new distribution is speedy and predictable.

Third, La Porta and others examine how the legal origin of a country affects who has power in the case of competing claims. Building on the classification provided by comparative law scholars, they focus on two legal traditions, common and civil law. In the common law tradition of the United States, the United Kingdom, and the Commonwealth countries, the law evolves as judges resolve specific disputes. In contrast, civil law relies on written statutes often based on abstract principles, with roots in Roman law. In civil law countries the formal nature of the law increases its predictability, but it also allows for more abuses of investors, as courts have demonstrated less willingness to involve themselves in disputes between insiders and outsiders when the transaction being evaluated might have a plausible business purpose.

Courts' lack of involvement in such disputes in civil law regimes stems from the interpretation of insiders' "duty of loyalty" to outside investors. As Johnson and others (2000) suggest, in common law countries the duty of loyalty clause has been seen as a residual concept that judges have used expansively, effectively leading to a lower standard of proof for inviting the courts to examine behavior that benefits insiders at the expense of outside investors. In countries with a civil law tradition, such as Belgium, France, Italy, and Japan, the same duty of loyalty clause has been interpreted narrowly and the courts have not gotten involved. Consequently, investors effectively have greater power in common law regimes.

BEYOND LEGAL RULES: COMPLEMENTARY INSTITUTIONS. The work of La Porta and others has stimulated interest in how institutional factors can affect investment behavior. The work clarifies the importance for governance of resolving competing claims on the wealth of the corporation and the effect of default rules and legal origin on the costs of resolving disputes. The research also has clear policy implications for how to direct reforms by showing that clarifying power and allocating power to providers of finance can foster investments.

Three types of criticism have been voiced against this new focus on legal rules. First, legal rules focus primarily on one function of governance systems—reducing the costs of resolving competing claims over the wealth of the firm. This is an important function but not the sole function of a governance system. Even with low costs in resolving disputes, there remain agency problems within firms as a result of inevitable information gaps between users and providers of information. Additional functions of governance systems are to reduce information gaps and manage the inevitable incentive problems stemming from these gaps that cannot be addressed through corporate laws. Addressing these functions of governance requires different institutions, including institutions that enhance penalties provided by product and labor markets.

The second and third criticisms suggest that legal rules do not necessarily lower the costs of resolving disputes. A view prominent in the law and economics literature is that the only legal protections required are contracts and a judicial apparatus to enforce contract law. National legal protections are neither required nor likely to be important. They are not required because sophisticated investors can negotiate complex contracts with sophisticated insiders in the firm to anticipate eventualities and build in contractual safeguards. And they may not be important because sophisticated investors can contract out of or expand on these protections. This criticism, while logically consistent, is not supported by the data I describe below. Legal rules do matter for the effectiveness of governance.

The third criticism, which I pursue here, does not take issue with the contention that rules matter but suggests that the impact of legal protections on corporate governance depends critically on other institutions in an economy. A simple reading of the message of La Porta and others can lead to misunderstanding. It is a mistake to interpret this work as calling solely for the adoption of specific laws. It is a further mistake to focus solely on the small number of provisions in company law that La

Porta and others suggest clearly allocate power to insiders or financiers. A constant, but often ignored, element of the authors' message is the importance of effective legal protections—both the rules and their enforcement.

A change in legal protections will not affect investor beliefs if the actors tasked with resolving competing claims and enforcing judgments do not have the incentive to do so. Nor will it shift power to investors if it is not accompanied by incentives to provide them with information so that they avail themselves of those protections.

There is ample evidence that legal protections alone are insufficient to improve governance. Russia's experience is extreme but not unique. Black, Kraakman, and Tarassova (1999) suggest that with the perfection of legal protections in the Russian Federation (many of which they helped draft), effective protections for investors declined:

[T]he principal problem is not that the laws aren't strong enough, but that they aren't enforced . . . unhappy shareholders can rarely develop enough facts to prove the rampant self-dealing that occurs every day. The courts respect only documentary evidence, which is rarely available, given limited discovery and managers' skill in covering their tracks . . . [P]ursuing a case . . . will take years, and when you're done, enforcing a judgment is problematic, because enforcement is by the same biased or corrupt lower court that the shareholder began at. (p. 1754)

What must complement legal rules to produce effective legal protections? Theory alone cannot provide the answer. But the history of institutions in developed markets provides a guide. Studies of that history suggest that the effectiveness of legal protections rests substantially on three sets of additional institutions that create incentives for enforcement and information provision: the organization of political authority to limit arbitrary actions of the sovereign, the presence of intermediaries that reduce the public good problem associated with monitoring insiders, and the provision of incentives for these intermediaries through appropriately designed regulation or product market competition.

JUDICIAL EFFICIENCY AND THE INCENTIVES OF THE SOVEREIGN. For legal protections to influence investment behavior, there must be expectations that those laws will be obeyed. Judges must be sufficiently skilled to identify the laws that apply and to interpret the evidence in light of those laws. And judicial actors must have an interest in enforcing the law. While the range of factors that create conditions for effective enforcement remains an open question, in the Western economies effective enforcement has been linked to political structures that limit the power of the sovereign relative to other interest groups in society.

This idea is often linked to the influential work of Douglass North, particularly North and Weingast's (1989) work on the role of structural reforms in the development of external finance in the United Kingdom. Rajan and Zingales (2000) present similar arguments, emphasizing how legal protections can be changed by a deter-

mined government, using evidence from 19th- and 20th-century Europe to bolster their arguments.

North and Weingast (1989) focus on a critical period of development for the United Kingdom at the end of the 17th century. Common law courts had been in existence for a long time, but their relevance waned in the 17th century under the Stuart monarchy, with the increasing role of prerogative courts, the centralization of power in the Star Chamber, and the use of the executive power of pay and appointment of judges to influence their decisions. To resurrect the role of common law, complex organizational reforms were undertaken—among other things, abolishing prerogative courts, introducing new legislation mandating lifetime judicial appointments (with removal from office possible only in the case of criminal behavior or the assent of both houses of Parliament), and imposing constraints on the sovereign (including the Bill of Rights). These changes were supported by both inducements for the sovereign (a fiscal revolution that provided a steadier stream of revenues) and credible threats against the sovereign (the successful replacement of two sovereigns over the previous 50 years). In short, what made the common law courts predictable venues for resolving commercial disputes was the creation and development of a complex set of self-supporting political institutions.

INSTITUTIONS THAT PROVIDE INFORMATION AND INCENTIVES FOR EFFICIENT INFORMATION PROVISION. The costs of providing the functions of governance depend in part on the ability of resource providers to bridge the information gaps between themselves and the insiders in firms. Financiers must have information if they are to avail themselves of the power granted them by legal protections. Such information also has independent value: the reduction of information asymmetries is one of the core functions of a governance system.

Some information is, of course, collected independently by investors and provided by the users of the resources. But on both the demand and the supply side the extent of information collection is limited without other intermediaries. On the demand side is the well-known public good problem associated with monitoring insiders. To the extent that there are costs in collecting and evaluating information, small investors are better off if they can rely on other investors to provide information. On the supply side the user of resources often faces a conflict of interest in revealing information about the use of those resources. The ability to rely on reputation mechanisms to enforce information provision is constrained by the fact that users of resources may ask for investments only intermittently and penalty mechanisms may be weak.

Intermediaries that certify the quality of information produced, analyze the information for evidence of diversion of resources from promised uses, and provide summary statistics to investors can play a critical role in governance. They can help alleviate the public good problem associated with monitoring management and lower the transactions costs of collecting information to improve overall information flows. No longer is information about reliability possessed only by those with

direct experience; specialized agencies now collect the information and make it available to clients.

Intertwined with the rise of trade credit for American retailers in the 19th century was the development of financial reporting, external auditors, and credit rating agencies (Olegario 1999). Organizations that provided information on assets and those with liens on the assets, such as property registries, helped facilitate credit and the use of collateral. The rapid expansion of railways and the demands for external finance through bond offerings increased demands for public accounting and encouraged the development of bond rating companies. Stock markets introduced listing requirements and brokerage houses, and the analysts that brokerage houses employed provided analyses of firms' current and future prospects.

Information provision depends on more than the existence of intermediaries, however. The expected efficiency of information intermediaries, which is what matters for investor confidence, depends on their access to information, their ability to process the information, and their incentives. A potential source of incentives is the prospect of competition among intermediaries or in labor markets. In Western economies there is evidence that the career prospects of securities analysts in brokerage houses are influenced by the accuracy of their earnings forecasts and their ranking by the institutional investors that use their reports (Hong, Kubik, and Solomon 2000). Brokerage houses whose analysts fail to provide accurate analyses lose market share.

Regulation of financial intermediaries appears to have been critical in most developed capital markets and in some nascent markets. To provide information and lower the costs of enforcement it may be optimal not to keep judicial authority in the general court system but to delegate authority to a specific regulatory agency—or even to allow that regulator to delegate its authority to private sector organizations. This is not an argument that delegation is always superior—giving authority to a specialized institution might lead to abuse—but that the location of authority can affect efficiency.

An analogy to agency problems within firms is appropriate. Firms do not completely centralize or decentralize decisionmaking. Instead, they identify who has the relevant information and combine provision of incentives with delegation of authority to those with information. The same logic applies one level up. A system of legal protections will be ineffective if it merely demands information or sanctions. A more effective system identifies those with potential access to information, delegates authority, and harnesses the incentives of the decisionmakers to whom authority is delegated.

Many examples illustrate the importance of the design of oversight institutions in ensuring that legal protections fulfill their desired functions—providing information and incentives for promise fulfillment and lowering the costs of resolving competing claims on the wealth of the firm. One of the best examples is the formation of the U.S. Securities and Exchange Commission (SEC) and its impact on markets for external finance, a system that facilitates more equity investment in corporations than any other such system in the world.⁷

The importance of high-quality information for investors had long been appreciated, but the stock market crash of 1929 revealed the inadequacy of previous attempts to address information asymmetries through simple disclosure rules. The faith of investors was restored by the reform of securities market regulation through the passage of the Securities Act of 1933, the Securities Exchange Act of 1934 (which created the SEC), and years of work by SEC officials to get the details right.

To reduce information asymmetries and increase incentives for disclosure and enforcement, the SEC imposed some penalties on firms issuing securities. Particularly important were penalties that had an effect immediately, even if subject to later judicial review.⁸ Just as important was the SEC's focus on critical third-party agencies, as McCraw (1982) emphasizes. By regulating those intermediaries and granting them significant powers, including the power to police themselves, the SEC used its informational advantages and lowered transactions costs for investors in resolving conflicts.

More recent evidence of the importance of legal protections and their enforcement through specific regulatory authorities has been offered by analysts comparing the Czech and Polish experiences in the transition toward a market economy (see, for example, Johnson and Shleifer 1999). The Czech Republic and Poland started their reforms with roughly similar (low) levels of investor protections, but they had dramatically different experiences with corporate governance. The Czechs suffered widespread looting of firms by managers and insiders and rapid loss of faith in the stock market, with delisting of firms and no new private companies. In Poland much less investor dissatisfaction has been reported, and many more new firms have been able to raise external capital. Poland's relative success has been attributed in part to its far more stringent regulation of securities, including the ability of the securities regulator to control financial intermediaries through licensing, the regulator's much greater requirements for disclosure by issuers of securities, and the *ex ante* restrictions that limit conflicts of interest for intermediaries. In contrast, there was little initial regulation in the Czech Republic, and few changes in the approach despite deteriorating economic performance.

The importance of appropriate regulation of financial intermediaries, whether through a government agency or a self-regulating organization, applies with equal force to credit markets. Central banks and bank regulators are particularly important, as they collect information on banks as intermediaries and can coordinate responses through their influence over banks. Their powers include the ability to withdraw licenses, impose capital adequacy requirements, restrict loans to classes of debtors, and even trigger bankruptcy proceedings for firms or banks.

Fulfilling the function of providing information requires the creation of appropriate incentives for intermediaries. Evidence from developed and nascent capital markets suggests the importance of a specialized agency and the value of allowing that agency to delegate authority to self-regulating organizations—where such organizations stand a chance of being effective. In the United States the focus on regulating intermediaries rather than the issuers of securities economized on

scarce regulatory resources. Moreover, the SEC harnessed the incentives of intermediaries, and SEC regulators by and large have not used their authority, as they might have, to reward themselves.

* * *

The argument so far can be summarized succinctly. Theory and historical evidence from some developed economies suggest that legal protections matter for corporate governance. But where legal protections are effective, they are accompanied by other institutions. These complementary institutions include elements of political organization, such as constitutions and decentralization, which shape the incentives of the sovereign and increase the power of judicial institutions. Complementary institutions also include information intermediaries and the regulatory organizations that provide intermediaries with the appropriate incentives.

Recognizing the extensive demands for institutional change and the fact that changes take place through a political process has important implications for the dynamics of governance reforms through legal channels. Changes in legal protections affecting financiers, the power of the sovereign, and the regulation of intermediaries are not Pareto improvements: insiders who benefit from the existing situation will lose as a result of reforms. Changes will be the outcome of a political process involving competition among interest groups. As a result, while legal reforms are important, they are unlikely to occur rapidly or to be right in the first instance. Such limitations should lead those interested in improving governance to look for other channels to increase the prospects for promise fulfillment.

Internal Governance Policies: Ownership Structure

An entirely different approach to addressing the functions of corporate governance is to look inside the firm and focus on policies at the discretion of those in control. Many policies—such as the structure and role of the board of directors, the design of executive compensation, and the financial structure—produce information, manage risks and incentives, and resolve competing claims. All of these are important, but they are beyond the scope of this article. Here I concentrate on ownership structure, both the identity and the concentration of owners. Are ownership structures that deviate from the dominant neoclassical model of the small, anonymous, diversified shareholder a second-best approach to provide the functions of governance when legal protections are ineffective?

The impact of ownership structure on the functions of governance is not nearly as straightforward as the impact of legal protections. Deviations from ownership by small anonymous shareholders have ambiguous effects. Since the early work of Berle and Means (1933), scholars have been concerned with the small management stake in large firms and the consequent misalignment of interests. Management ownership—at least at some levels, as Morck, Shleifer, and Vishny (1986) argue—has been seen as a response that improves firms' performance. But Demsetz and

Lehn (1985) and others highlighting the endogenous determination of ownership stakes when the transactions costs of reallocating ownership are low have found no effect on performance.

The main argument advanced below sidesteps this debate by looking across countries where there are significant differences in the effectiveness of legal protections for investors. The contention is that where legal protections are weak as a result of institutional weaknesses, ownership structure has systematic effects, both in theory and in practice. Like legal protections, ownership structure can provide the functions of corporate governance.

PROMISE FULFILLMENT THROUGH OWNERSHIP IDENTITY. Consider a situation with no effective legal protections. In this circumstance investors are likely to be extremely reluctant to give up resources to someone else in exchange for a promise, because if that promise is violated there is no clear penalty they can impose. There is thus the danger of no investment in corporations. The theory of repeated games suggests one way out of this trap. The prospect of linking current violations of promises to future penalties can lead those in control of investor resources to honor their promises. But bilateral reputation mechanisms will often not achieve an efficient level of investment. If insiders in the firm violate the interests of one financier, they may lose the ability to appeal to that financier again, but this threat is a weak deterrent if they can freely access finance from other financiers. Reputation mechanisms can be made more effective if a way is found to magnify the penalty for failing to fulfill promises—in effect replacing bilateral reputation with multilateral reputation.

Where networks exist that provide information, allow for coordinated action, and can enforce such action, the identification of an individual who receives an investment with a network can create a multilateral reputation mechanism. The state's enforcement of legal protections is the ultimate example of such a multilateral mechanism. But it is not the only one. Greif (1997) shows that augmented reputation feedback mechanisms helped drive the commercial revolution—such mechanisms as the community responsibility system, in which any member of a community could be held liable for unmet promises of another member; the identification of traders with specific religious communities, such as the Maghribi; and even artificial communities created by merchants, such as the German Hansa. In all these examples the identification of an individual with a larger community helped increase investment opportunities.

When the controller of investor resources is not anonymous, it is possible to go from no investment to investment. Community affiliation conveys information about the penalties that will be imposed if the controller of the investment violates the terms of agreements—and thus affects investors' beliefs. This deceptively simple logic, constructed to explain historical evidence on the emergence of long-distance trade, has much broader scope. It helps explain why some contemporary economic agents are able to support investments while others have difficulty doing so.

The family-centered and ethnically based business groups common to much of the developing world are networks in which identity matters. Members of the business group have reputations, can collect more information than individuals who are

not members, and can collectively enforce penalties. In theory, these advantages facilitate investments that would not be made in the absence of these community ties. Summarizing the evidence on the impact of group affiliation, Khanna (2000) finds that group affiliation often helps rather than harms performance. This result is consistent with the framing described above, but it is also consistent with other theories of group formation.

Business associations are artificial communities that can also support investments in firms if they collect information, coordinate action, and have the incentive to enforce penalties for noncompliance. There is recent evidence of their ability to support trade credit in the weak legal environments of Poland, Romania, Russia, and Ukraine (Johnson, McMillan, and Woodruff 1999) and in that of Vietnam (McMillan and Woodruff 1998).

Foreign owners can also be considered members of networks, which might partly explain the superior performance of foreign-controlled firms in developing markets in India (Chhibber and Majumdar 1999) and the transition economies (Frydman and others 1999; Djankov 1999; and Djankov and Murrell forthcoming). When foreign owners assemble a controlling stake, for example, they become subject to their home country's regulations (both legal requirements and any additional requirements by the stock exchange that lists their shares). These regulations often require them to disclose far more information than demanded by the country in which the investment is located, and the penalties for violating these rules are often the same as if the investment were in the home country. Penalties can also be applied to assets in the firms' home country. Thus foreign owners post something greater than their word when operating in a foreign country.

The argument that identity supports investments is not an argument that identity should be used rather than legal protections; relative to the first-best solution, identity has clear limitations. The ability to pool investments, allocate them across space, and manage risks is constrained by the need for those controlling investments to have clear ties to an identifiable community. There may be inefficiently low levels of information collection by investors.⁹ The costs can be high, producing inflexibility. Members of a community that collects private benefits from the networks' presence have a strong interest in seeing that these mechanisms are maintained, even if they are no longer the most efficient mechanism.

Such systems are also likely to introduce added costs when the economy faces temporary or permanent shocks. As Johnson, McMillan, and Woodruff (1999) find, investors relying on informal mechanisms are inclined to maintain those relationships rather than use other mechanisms, even when there are costs to doing so. The costs of relying on reputation mechanisms to support promised returns to investments are unusually sensitive to the future growth prospects of the firm and the need to turn again to outside finance. When the very future of the firm is in question as a result of a change in economic circumstances, such as an economic crisis, even "reputable" owners might engage in unrestrained looting of the firm.

PROMISE FULFILLMENT THROUGH OWNERSHIP CONCENTRATION. A second deviation from anonymous, dispersed shareholding that can enhance promise fulfillment is ownership concentration. The traditional motivation for concentration is that it reduces the public good problem associated with monitoring insiders. The greater the ownership stake, the greater the personal returns to monitoring insiders and exercising voice and the more information owners will collect. Coffee (1999), among others, suggests that the activism of shareholders is proportional to the concentration of ownership. Concentrated owners often become what Jensen (1991) calls the active investor: "one who holds large equity and/or debt positions and actually monitors management, sits on boards, is sometimes involved in dismissing management, is often closely involved in the strategic direction of the company and, on occasion, even manages" (p. 15).

By monitoring managers, concentrated shareholders might also enhance information flows for other providers of resources, particularly if there is a forum to share this information (say, a board of directors). This channel for reducing information gaps is particularly important when information provided through public channels is opaque and not timely.

Bebchuk (1999) offers a more recent rationale for concentration in the presence of ineffective legal protections. Bebchuk focuses on how concentration can resolve competing claims on the wealth generated by the corporation. He suggests that concentrated ownership is the only viable structure in countries with weak legal protections. Where legal protections are weak, the scope for redirection of resources is large. Faced with this situation, a single owner wanting to sell a stake would be much better off selling a large controlling stake than selling dispersed shares. The problem with dispersion is that each dispersed shareholder anticipates that someone will assemble a controlling stake and use that control to redirect resources to himself or herself. Concern about this eventual theft will lower share prices today. The single shareholder maximizes his or her returns to sale by maintaining control; control protects rents.

Concentration may also enhance efficiency. With dispersed shareholding and possibilities for diverting returns, it may not be clear initially which of the dispersed shareholders will assemble a controlling stake. In such a situation those with limited control today but uncertainty about future control have an incentive to dilute quickly and destructively. In contrast, if control is obtained by having a large cash flow stake in the firm, those in control know they will retain control and their temptation to dilute is limited by their ownership stake. At least to some extent, they are simply stealing from themselves. Although concentrated control may not be the only viable option in countries with weak legal protections, it may be more efficient than dispersed shareholding.

In environments with weak legal protections, it may be optimal to have more than one large owner, particularly if the new owners have identifiable characteristics tied to the functions of corporate governance. For example, a potential investor with financial capital might be more willing to invest if suppliers or buyers also have stakes, because even without legal protections these agents have potential leverage over the insider (hold-up power). The small investor knows that if insiders attempt

to reallocate promised returns, financiers and suppliers as cofinanciers can credibly threaten to penalize them directly, thereby reducing the likelihood of such an event. Aoki (1984) underscores how labor might be able to police against abuse through its ability to withhold services. Financial institutions are well positioned to play this role. Their hold-up power derives from their ability to cut off not only long-term finance but also short-term capital.

Through the early 1990s many analysts of corporate governance systems in Germany and Japan suggested that a key to the relatively strong performance of firms in these countries was the ownership and control structure, which involved banks, workers, and concentrated owners. These parties have information, the ability to impose sanctions, and the incentive to use their powers. Investors believed that the presence of a bank mattered, and this affected their willingness to invest. Kaplan (1994) and Kaplan and Minton (1994) show that there is little difference in the sensitivity of management turnover to changes in financial performance between Germany and Japan, which rely more heavily on informal mechanisms, and the United States, with its largely formal system. Hoshi, Kashyap, and Scharfstein (1990) show that Japanese firms with bank relationships have greater access to capital and lower costs of resolving financial distress than firms without bank relationships.

As with ownership identity, the argument here is not that ownership concentration is superior to a governance system with effective legal protections. Instead, the argument is that ownership concentration can provide the functions of governance and that, particularly where legal protections are weak, the benefits provided by concentration outweigh the costs.

In fact, the costs can be sizable, for several reasons. First, with concentrated ownership, there is no longer separation and specialization based on the comparative advantages of those providing management and those with capital to invest. Risks are managed less efficiently, and the pool of investors is more limited. Second, as Bolton and von Thadden (1998) emphasize, there is a loss in liquidity, with a resulting decline in the value of shares. Third, with concentrated ownership, those in control are difficult to dislodge. When these insiders are not well motivated, efficiency will decline. Aghion and Blanchard (1996) emphasize the problems that arise when workers or managers have control of the firm. Recent literature on the German and Japanese systems now sees bank affiliation as also costly, with bank involvement leading to overlending and deferred restructuring (Weinstein and Yafeh 1998). Fourth, to the extent that initial owners maintain their concentration through voting rights rather than cash flow rights, incentives to use control in the interests of minority shareholders are reduced. There are various vehicles through which owners can maintain control disproportionate to their cash flow stake—shares with different voting rights, cross-shareholding, or pyramid structures, for example. As discussed below, Claessens, Djankov, and Lang (2000), in their study of Asian firms, provide evidence of costs associated with concentration of control when it is not linked with similar cash flow rights.

* * *

Governance policies, like institutions, clearly provide the functions of corporate governance. Relative to anonymous, dispersed shareholders, owners who are identified with a network and have significant concentration can have a comparative advantage in providing information, managing incentives, and lowering the costs of resolving competing claims on the wealth of the firm. The evidence presented here comes mainly (but not exclusively) from countries in which legal protections are weak. This is consistent with theory, which shows that with weak legal protections other ownership structures are costly and may even be unsupported. When legal protections are strong, the theoretical results are more ambiguous, as is the evidence.

International Data on Legal Protections and Ownership Structure

The previous section provided a conceptual framework and some suggestive evidence on how to think about institutions and policies that can provide the functions of governance. More systematic studies of the relationship among governance institutions, governance policies, and firms' performance extend these arguments. I first examine cross-sectional data for established firms and then look at data for privatized firms.

Established Firms

Among the most notable studies producing comparable cross-country data on governance institutions and policies are La Porta and others (1998); La Porta, López-de-Silanes, and Shleifer (1999); Pistor, Raiser, and Gelfer (2000); Claessens, Djankov, and Lang (2000); and Becht and Roell (1999).

La Porta and others (1998) collected information on corporate governance characteristics and cash flow ownership for a cross-section of 49 countries. Their sample is heavily weighted with higher-income economies—it covers 93 percent of high-income economies and 62 percent of upper-middle-income economies—but it still covers a significant number of developing countries, with a third of the sample defined by the World Bank as low or middle income.¹⁰ Pistor, Raiser, and Gelfer (2000) provide data on legal protections for 24 transition economies from 1992, when the transition began, through 1998.

La Porta, López-de-Silanes, and Shleifer (1999) provide an alternative measure of ownership concentration for a sample of 691 firms in the 27 most developed countries. This study has been complemented by more detailed studies of the concentration of voting rights in particular regions. Becht and Roell (1999) report on studies based on 1,381 firms in eight European countries under the umbrella of the European Corporate Governance Network. Claessens, Djankov, and Lang (2000) study 2,980 firms in nine East Asian economies.

LEGAL PROTECTIONS FOR EQUITY INVESTORS. The data reveal surprising heterogeneity in investor protections around the world—and a significant number of countries with weak protections. To illustrate these differences, I focus on protections for equity investors—the six anti-director rights (see the section on national legal pro-

tections). Of a maximum score of 6 for anti-director rights, the highest score in the country sample in La Porta and others (1998) is 5 and the average just 3 (table 1). Interestingly, the scores for anti-director rights are not driven by differences in per capita income, with average levels indistinguishable at different income quartiles.

Pistor, Raiser, and Gelfer (2000) evidence for the transition economies expands our understanding of legal protections in countries not classified as high income. Echoing the findings of La Porta and others (1998), they find wide variation in the extent of legal protections across countries and many countries with relatively weak protections. They also document improvement in the transition economies, where the average score for anti-director rights rose from 1.8 in 1992 to 3.0 in 1998—identical to the average for the established economies investigated by La Porta and others.

The change in the extent of legal protections from 1992 to 1998 in the Pistor and others sample demonstrates one reason why legal protections have become the subject of such policy interest. Unlike the complementary institutions described above, legal protections can be changed very rapidly. And efforts by the international community matter. As Pistor, Raiser, and Gelfer (2000) show, countries that received U.S. aid have more legal protections than other countries in the sample, including those preparing for membership in the European Union.

COMPLEMENTARY INSTITUTIONS. Statistical research has not focused on the complementary institutions of judicial efficiency, information intermediaries, and the circumstances that create incentives for those intermediaries. Where these complementary institutions have found their way into the analysis, they have done so through relatively crude proxies. A number of studies use “rule of law” or “efficiency of the judiciary” indexes compiled by commercial risk agencies as proxy measures for judicial efficiency. To capture the presence of publicly available information, these studies use an index that indicates whether 90 factors identified by accountants as useful indicators of a firm’s financial affairs are included in the company’s annual report. (See La Porta and others 1998 for a description of these variables and their sources.)

The extent of the rule of law is very closely linked with per capita income: the correlation coefficient between the log of per capita GDP and the rule of law measure is 0.87. The accounting standards index ranges from a low of 24 (out of 90) for Egypt to a high of 83 for Sweden (data are unavailable for some countries). The correlation coefficient with respect to the log of per capita GDP is a lower but still significant 0.51.

While sometimes criticized as an artificial distinction by legal scholars, legal origin has surprising explanatory power, significantly affecting anti-director rights and other variables. The strongest difference is between countries with a French civil law origin and those with a common law origin. The mean index for anti-director rights, the rule of law, and accounting standards is higher in common law countries, and the difference is significant for anti-director rights.

For the purposes of presentation in this article, I create an “effective legal protections index” that more directly captures the conjectured complementarity of legal

Table 1. Legal Protections in Selected Economies

<i>La Porta and others (1998) sample</i>					<i>Pistor and others (2000) sample</i>				
<i>Country and legal origin</i>	<i>Anti-director rights index (6)</i>	<i>Rule of law (10.00)</i>	<i>Accounting standards index (90)</i>	<i>Effective legal protections index (100)</i>	<i>Country</i>	<i>Anti-director rights 1992 (6.00)</i>	<i>Anti-director rights 1998 (6.00)</i>	<i>Rule of law index (10.00)</i>	<i>Effective legal protections index at beginning of privatization (100)</i>
English common law									
Australia	4	10.00	75	67	Armenia	2.50	5.50	4.9	20
Canada	5	10.00	74	83	Azerbaijan	2.50	2.00	3.2	11
Hong Kong	5	8.22	69	69	Belarus	1.50	1.50	2.3	6
India	5	4.17	57	35	Bulgaria	4.00	4.00	5.9	39
Ireland	4	8.75	—	52	Czech Republic	2.00	3.00	8.3	28
Israel	3	4.82	64	24	Estonia	2.00	3.75	8.5	28
Kenya	3	5.42	—	27	Georgia	2.50	3.00	4.0	17
Malaysia	4	6.78	76	45	Hungary	2.50	3.00	8.7	36
New Zealand	4	10.00	70	67	Kazakhstan	2.50	5.25	4.5	19
Nigeria	3	2.73	59	14	Kyrgyz Republic	2.50	2.25	4.4	18
Pakistan	5	3.03	—	25	Latvia	2.50	3.50	7.5	44
Singapore	4	8.57	78	57	Lithuania	2.50	3.75	7.2	30
South Africa	5	4.42	70	37	Moldova	3.00	3.50	4.7	24
Sri Lanka	3	1.90	—	10	Poland	3.00	3.00	8.7	44
Thailand	2	6.25	64	21	Romania	3.00	3.00	5.6	28
United Kingdom	5	8.57	78	71	Russian Federation	2.00	5.50	3.7	12
United States	5	10.00	71	83	Slovak Republic	2.50	2.50	6.4	27
Zimbabwe	3	3.68	—	18	Slovenia	0.00	2.50	8.4	0
French civil law									
Argentina	4	5.35	45	36	Ukraine	2.50	2.50	3.4	14
Belgium	0	10.00	61	0	Uzbekistan	2.50	3.50	2.7	11
Brazil	3	6.32	54	32	Average	2.44	3.00	5.82	22
Chile	5	7.02	52	59					
Colombia	3	2.08	50	10					
Ecuador	2	6.67	—	22					
Egypt, Arab Republic	2	4.17	24	14					
France	3	8.98	69	45					
Greece	2	6.18	55	21					
Indonesia	2	3.98	—	13					
Italy	1	8.33	62	14					
Mexico	1	5.35	60	9					
Netherlands	2	10.00	64	33					
Peru	3	2.50	38	13					
Philippines	3	2.73	65	14					
Portugal	3	8.68	36	43					
Spain	4	7.80	64	52					
Turkey	2	5.18	51	17					
Uruguay	2	5.00	31	17					
Venezuela, Republica Bolivariana de	1	6.37	40	11					
German civil law									
Austria	2	10.00	54	33					
Germany	1	9.23	62	15					
Japan	4	8.98	65	60					
Korea, Republic of	2	5.35	62	18					
Switzerland	2	10.00	68	33					
Taiwan, China	3	8.52	65	43					
Scandinavian civil law									
Denmark	2	10.00	62	33					
Finland	3	10.00	77	50					
Norway	4	10.00	74	67					
Sweden	3	10.00	83	50					
Average	3.0	6.9	60.9	35.0					

— Not available.

Note: The numbers in parentheses are the maximum values for the indexes.

Source: La Porta and others 1998; Pistor and others 2000; and author's calculations.

protections and their enforcement (see table 1). This index is constructed by taking the simple product of the index of anti-director rights (rescaled to vary from 0-10) and the rule of law measure.

OWNERSHIP STRUCTURE. Who controls firms? Recent research has added greatly to our understanding of the types of ownership structures around the world (results from three of the most cited studies in this literature are reproduced in table 2). La Porta and others (1998) collected data on cash flow ownership for the same sample of 49 countries described above. But cash flow ownership might not capture who controls firms when voting concentration differs from cash flow concentration. To determine whether any entity has sufficient voting rights to control a firm, La Porta, López-de-Silanes, and Shleifer (1999) track ultimate owners through pyramids and cross-shareholding arrangements. A shareholder with a direct or indirect voting stake of 20 percent or more is said to have control.

So what does ownership look like? The strongest finding from all these studies is that around the world, firms held by anonymous, dispersed shareholders are the exception rather than the rule. This is clearly demonstrated in the first column in table 2, which shows that the combined stakes of firms' three largest shareholders (based on cash flow) average a surprisingly high 46 percent. Only in Japan, Taiwan (China), the United Kingdom, and the United States do the combined stakes average 20 percent or less.

Using voting rights to measure concentration reinforces the conclusion that dispersed ownership is rare. With control defined as a direct or indirect voting stake of 20 percent or more, only 36 percent of the largest firms are widely held in countries in the La Porta and others (1998) sample. These results are robust. The limited role of dispersed shareholders is evident in Europe: the data in Becht and Roell (1999) show that only 22.4 percent of firms in Austria, Germany, and the Netherlands have a controlling shareholder with less than a 25 percent stake. This result is echoed in Asia, where only 32 percent of firms are widely held (Claessens, Djankov, and Lang 2000). What is true for the largest firms around the world holds with even greater strength for medium-size firms. In the La Porta, López-de-Silanes, and Shleifer (1999) voting rights sample, for example, just 24 percent of medium-size firms are widely held. Within the sample, the smaller the firm, the greater the likelihood that there is a controlling shareholder.

The distinction between cash flow rights and voting rights turns out to be important, for the two are routinely separated. The evidence shows that the most common way to separate these rights is through the use of pyramid structures. Less common is cross-shareholding, although this is important in such countries as Japan. Surprisingly rare are instances of issuing multiple shares with different voting rights.

Jensen's (1991) contention, based on evidence from the United States, that concentrated owners often become active investors is consistent with the international evidence. In most firms with a controlling shareholder the distinction between owners and managers is eliminated. When a family is the controlling shareholder, family members participate directly in management in 69 percent of cases in the La

Table 2. Ownership Concentration As Measured by Cash Flow and Voting Stakes
(percent)

Country	Average combined cash flow stakes of firms' three largest shareholders ^a	Share of large firms with a controlling shareholder ^b	Share of medium- size firms with a controlling shareholder ^b
Argentina	53	100	100
Australia	28	35	70
Austria	58	95	100
Belgium	54	95	80
Brazil	57	—	—
Canada	40	40	40
Chile	45	—	0
Colombia	63	—	—
Denmark	45	60	70
Egypt, Arab Republic	62	—	—
Finland	37	65	80
France	34	40	100
Germany	48	50	90
Greece	67	90	100
Hong Kong	54	90	100
India	40	—	—
Indonesia	58	95	94
Ireland	39	35	37
Israel	51	95	90
Italy	58	80	100
Japan	18	10	70
Korea, Republic of	23	45	70
Malaysia	54	95	88
Mexico	64	100	100
Netherlands	39	70	90
New Zealand	48	70	43
Nigeria	40	—	—
Norway	36	75	80
Pakistan	37	—	—
Peru	56	—	—
Philippines	57	95	84
Portugal	52	90	100
Singapore	49	85	60
South Africa	52	—	—
Spain	51	65	100
Sri Lanka	60	—	—
Sweden	28	75	90
Switzerland	41	40	50
Taiwan, China	18	85	64
Thailand	47	95	94
Turkey	59	—	—
United Kingdom	19	0	40
United States	20	20	10
Venezuela, Republic Bolivariana de	51	—	—
Zimbabwe	55	—	—
Average	45.89	68.13	75.27

— Not available.

a. Data are from La Porta and others (1998). Large firms are defined here as the ten largest publicly traded companies.

b. A controlling shareholder is defined as one who has a direct or indirect voting status of 20 percent or more. Data are from La Porta, Lopez-de-Silanes, and Shleifer (1999) and Claessens, Djankov, and Lang (2000). Where results from the two studies differ, those from La Porta, Lopez-de-Silanes, and Shleifer (1999) are used. Large firms are defined as the twenty largest publicly-traded companies in 1995. Medium-sized firms defined as the twenty smallest publicly-traded firms with market capitalization greater than \$500 million in 1995.

Source: As specified in the notes above.

Porta and others sample and 67 percent in the Claessens, Djankov, and Lang sample for East Asia.

These data also allow an examination of the theoretical argument that multiple large shareholders may be important to facilitate investment in firms. For industrial economies there is little evidence of multiple large investors. When a family is in control, other concentrated owners are found in only 25 percent of cases in the La Porta, López-de-Silanes, and Shleifer (1999) sample. And in Europe the median stake of the second-largest shareholder is almost always below the minimum reporting level of 10 percent.

For developing economies there are fewer data, and the conclusions that can be drawn are less clear-cut. Interestingly, a second large shareholder is far more common in the Claessens, Djankov, and Lang (2000) sample of East Asian economies. In this more economically diverse sample, 49 percent of firms had a second shareholder with a significant stake. This evidence reopens the question of whether multiple large shareholders are important to governance in developing economies.

Privatized Firms

Privatized firms provide another data set that can be examined to study whether legal protections help explain ownership structure and firms' performance. In many ways privatized firms are ideal. They are often large firms, for which governance concerns are most important. The initial ownership at the time of privatization is in many ways a choice variable for politicians, so examining initial structures reveals whether political decisionmakers believe that the existing state of legal protections imposes some constraints on their actions. Most important, this is a good data set for examining whether the effectiveness of legal protections affects the evolution of ownership structures and firms' performance. If legal protections are a binding constraint, this should be reflected in disappointing performance and instability in environments with weak protections when ownership is dispersed. In these same environments concentration and identity should contribute to better performance.

No study provides the same systematic information on ownership concentration for privatized firms that has been collected for established, publicly traded firms. As an (imperfect) proxy, for which there are data, studies by Megginson and others (forthcoming) examine the relative tendency of governments to use asset sales rather than share issues to privatize state-owned enterprises in countries with established private sectors. Share issue privatizations are likely to produce ownership structures in which there is no controlling shareholder. Asset sales, in contrast, are usually associated with sales of a majority stake to a single investor or a consortium of investors arranged before the sale. Governments often establish prequalification criteria before the sale and conduct the sale through an auction or by direct sale to a targeted investor. Investigation of winning consortia shows that almost all have a core investor. Thus asset sales are likely to produce a controlling shareholder.

The transition economies represent a particularly rich source of data on legal protections and ownership structures. It is possible to construct a similar “effective legal protection index” as used for established firms. The effective legal protection index is based on the Pistor, Raiser, and Gelfer (2000) minority shareholder protection index in the year that marked the beginning of the country’s privatization program. The rule of law index is the score for that index in 1998, the first year with comparable information. Overall scores for effective legal protections are low. To repeat the obvious, the odds were against effective formal protections. The historical legacy of socialism included an atrophying of legal frameworks to support private ownership. Courts were not independent. Legislation was at times slow to develop, and in other cases what was written was weakly enforced.

While no single data source has collected data on the concentration of ownership in privatized firms, there is a high correlation between the privatization approach taken and the resulting ownership structure. The European Bank for Reconstruction and Development (EBRD 1999) classifies economies on the basis of whether the primary approach was voucher privatization, direct asset sales, or management and employee buyouts. A voucher privatization is a share issue privatization that results initially in dispersed ownership. Direct asset sales result in more concentrated ownership. Management and employee buyouts are somewhere in between voucher privatizations and direct sales in terms of creating concentration at the time of privatization.

Individual country studies provide better measures of ownership structures created through privatization programs. Djankov and Murrell (forthcoming) identify 23 studies that relate ownership structure to performance. Collectively, these studies introduce 11 different ownership structures, with a fair degree of overlap in classifications. The most common distinction is between private and state ownership, but private ownership is broken down into several categories, including dispersed ownership and concentrated ownership by type of owner. An important distinction is between firms in which the dominant shareholders are insiders and those in which they are outsiders. Another distinction is between types of outsiders, including banks, investment funds, and foreign owners.

Findings

It is possible only to scratch the surface of work that relates legal protections, ownership structure, and performance. Focusing on the impact of legal protections on ownership structure, financial performance, and investment, I first examine the data for established, publicly traded firms and then turn to evidence for privatized firms, which provides a more detailed picture of the relationship between ownership structure and performance.

Legal Protections and Ownership Structure in Established Firms

The extent of legal protections is highly correlated with ownership structure: the relationship between legal protections and ownership concentration for large firms

is strongly negative, as reflected in the downward sloping line of best fit in figure 1. Analysis by La Porta and others (1998) that controls for other contributors to concentration bears out these results. Such evidence is consistent with the contention that in established, publicly traded firms ownership concentration is a substitute for legal protections in providing the functions of corporate governance.

While anti-director rights are important, other types of legal protections also matter, as does their interaction. For the raw data, the explanatory power provided by anti-director rights and the rule of law is comparable. Anti-director rights explain 15 percent of the observed variation in the concentration for cash flow rights and 13 percent for voting rights; the rule of law explains 21 percent of the variation in the concentration of cash flow rights and 17 percent of that for voting rights. The effective legal protections index, that is the simple product of the anti-director rights and rule of law indices explains 32 percent of the variation in the concentration of cash flow rights and 28 percent of that for voting rights (figure 2). The point here is not to offer a new evaluation of the data but to reveal the empirical support for the contention that complementary governance institutions are critical contributors to the observed relationships between legal protections and other variables.

A second result, apparent in the plots of raw data, is that supporting anonymous, dispersed shareholders is particularly difficult in countries with weak legal protections, as shown by the blank space in the bottom left-hand corner in all the panels in figure 1 and figure 2. This finding is consistent with theory, which suggests that dispersed shareholding is unlikely to emerge and will not be sustainable with weak legal protections. The raw data suggest that the rule of law is particularly important. When concentration is measured by voting rights and the rule of law measure falls below 7, more than 10 percent of firms are widely held in only one country (the Republic of Korea).

A third result is that with effective legal protections it is hard to discern a distinct relationship between ownership and the quality of laws. While strong legal protections are necessary to support small, diversified shareholders, many countries that have such protections still use concentrated ownership. This is particularly true in Europe. This puzzle suggests either that the costs of ownership concentration have been exaggerated, or that other forces are at work to limit the ability to disperse ownership.

Legal Protections, Financial Performance, and Investment in Established Firms

Some of the most striking results relate institutions of corporate governance to financial performance. La Porta and others (2000) report that firms in countries with stronger legal protections are more likely to distribute earnings through dividends. Consistent with this result and with theory, La Porta and others (1999a) and Claessens, Djankov, and Lang (2000) show that legal protections enhance the value of shares (measured using Tobin's Q). The interpretation is that because legal protections convey power to minority shareholders, the market places a higher value on shares. As in studies relating legal protections to ownership, the impact of legal protections comes not only through anti-director rights but also through other variables.

Figure 1. Legal Protections and Ownership Concentration in Large Firms in Selected Economies

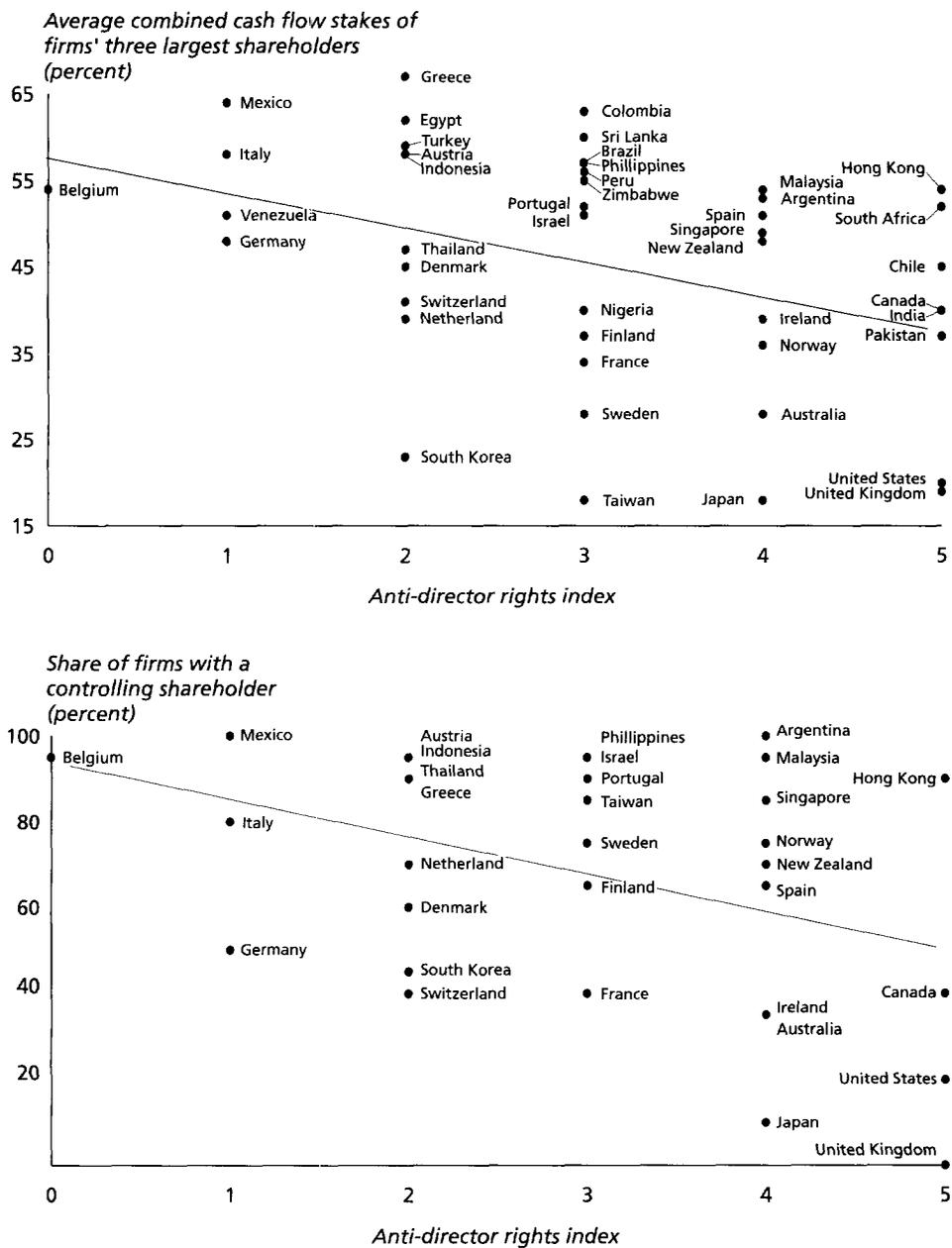
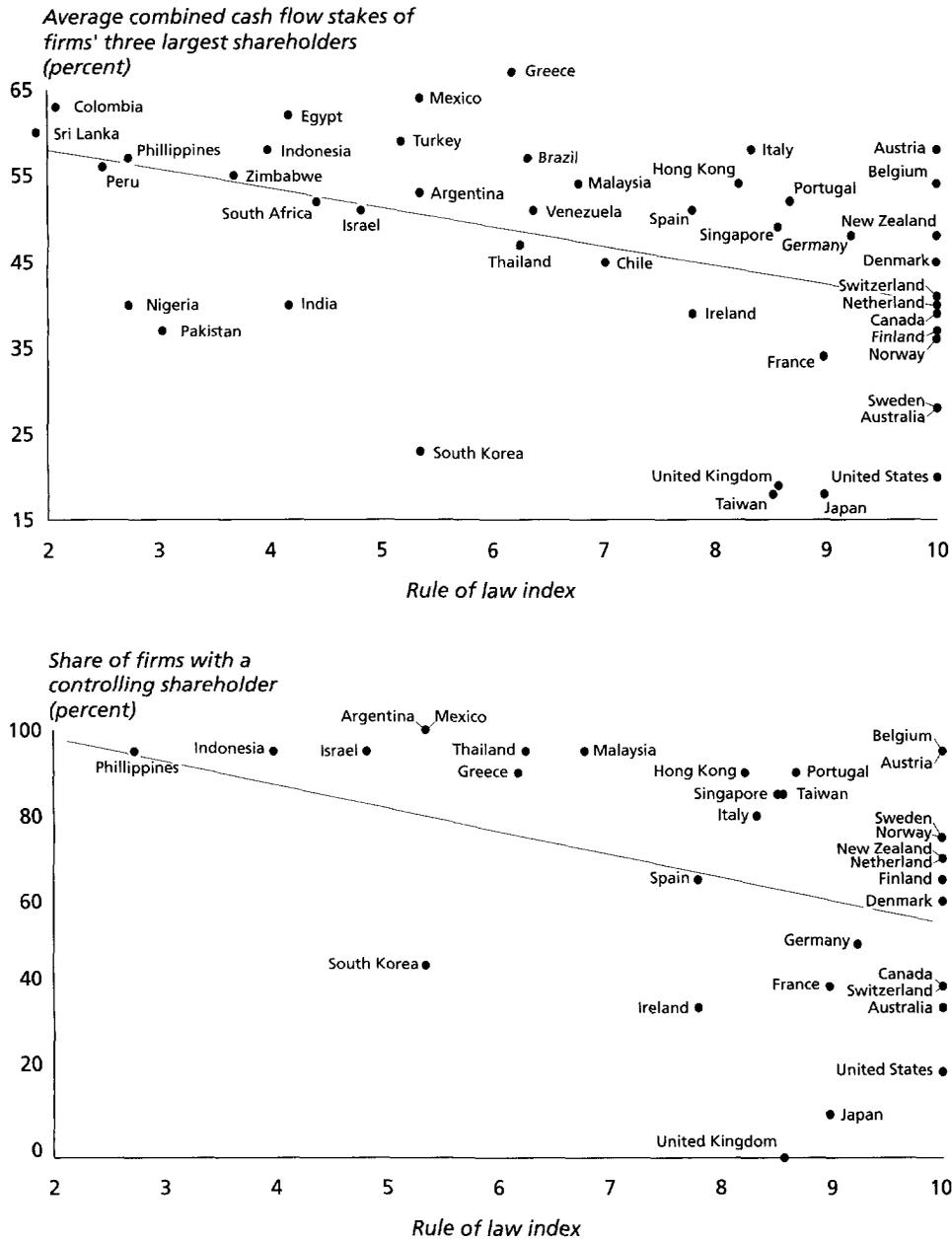
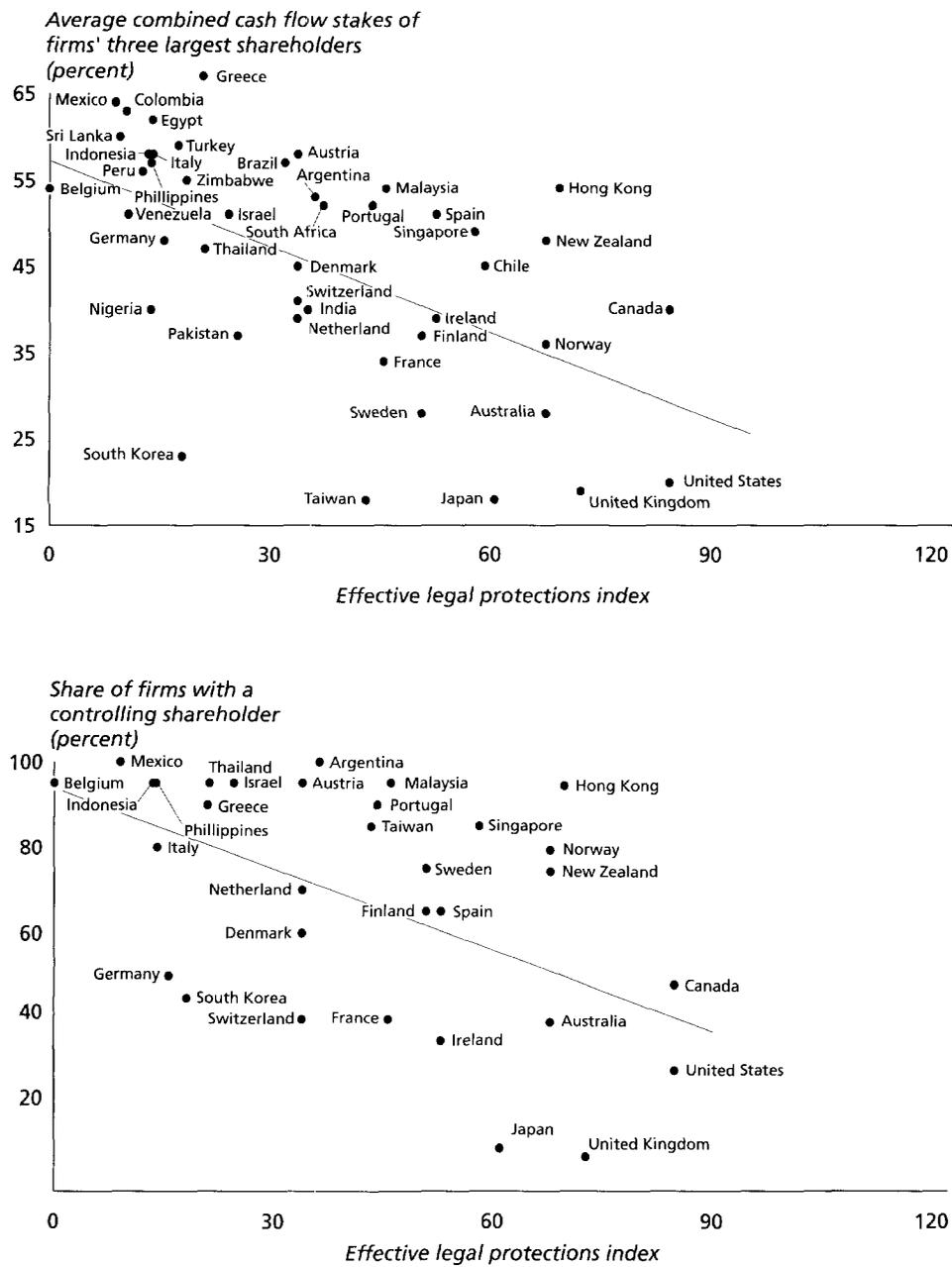


Figure 1. cont.



Source: For ownership concentration, La Porta and others (1998) and Claessens, Djankov, and Lang (1999); for legal protections, La Porta and others (1998).

Figure 2. Effective Legal Protections and Ownership Concentration in Large Firms in Selected Economies



Source: For ownership concentration, La Porta and others (1998) and Claessens, Djankov, and Lang (1999); for legal protections, author's calculations based on data from La Porta and others (1998).

Legal origins and, to a lesser extent, anti-director rights raise the value of shares.

These studies also show that ownership concentration through cash flow stakes is valued. To avoid confounding concentration with control, they look at concentration in firms that are “controlled” according to their measure of direct and indirect control. Consistent with theory, Claessens, Djankov, and Lang (2000) find strong evidence that concentrated cash flow ownership increases value, while concentrated voting rights reduce it. La Porta and others (1999a) present weaker evidence of the benefits of concentrated cash flow ownership.

Johnson and others (2000) illustrate that legal protections may be particularly important during times of crisis. They show that in East Asia during the recent financial crisis, countries with weaker formal protection mechanisms suffered larger declines in share prices and currency values than countries with stronger protections. These institutional variables have at least as much predictive power as standard macroeconomic variables.

Probably the most interesting findings are those that get directly at the aggregate measure of the effectiveness of governance—the efficiency with which resources are mobilized and reallocated within firms and across industries. La Porta and others (1997) provide evidence that the extent of legal protections for financiers is correlated with the depth of equity markets and the rate of initial public offering activity. The rationale is straightforward. With effective legal protections, investors have good reasons to expect that their promises will be fulfilled. Because they are more willing to exchange their resources for promises, the cost of financing drops and financial markets develop. Wurgler (2000) shows that the efficiency of the allocation of investment is related to legal protections. His estimates reveal a greater sensitivity of investment to growth opportunities in countries with more developed financial markets (an indirect link to legal protections) and a greater willingness to reduce investment in declining industries in countries with weak legal protections (measured as effective legal protections, in this case as the product of rule of law and the sum of creditor and anti-director rights).

The links between legal protections, firms’ performance, and economic growth remain indirect but very suggestive. Studies such as Rajan and Zingales (1998) show strong links between the extent of financial development and subsequent economic growth and development.

Legal Protections and Ownership Structure in Privatized Firms

The evidence for privatized firms reinforces these findings and adds support to the contention that ownership structure is important in countries with weak legal protections. In regressions that control for a variety of factors in privatizations in 80 countries, Megginson and others (forthcoming) report that legal protections and the government’s ability to credibly commit to property rights are both significant in explaining the choice of privatization approach.

The United Kingdom, for example, with its excellent legal protections, relied on share issue privatizations with widely dispersed shareholding. The average ownership stake of the largest shareholder for a sample of 25 electricity and water supply compa-

nies privatized in the United Kingdom was just 4.6 percent in the year of privatization (see Cragg and Dyck 1999b). In contrast, middle-income countries with weak effective legal protections used asset sales, which Megginson and others report accounted for 89 percent of privatizations in Argentina, 91 percent in Mexico, and 100 percent in Bolivia. In López-de-Silanes's (1997) privatization sample for Mexico controlling stakes were sold in 87 percent of firms; when noncontrolling stakes were sold, shares were purchased by the preexisting controlling shareholder in 83 percent of the cases.

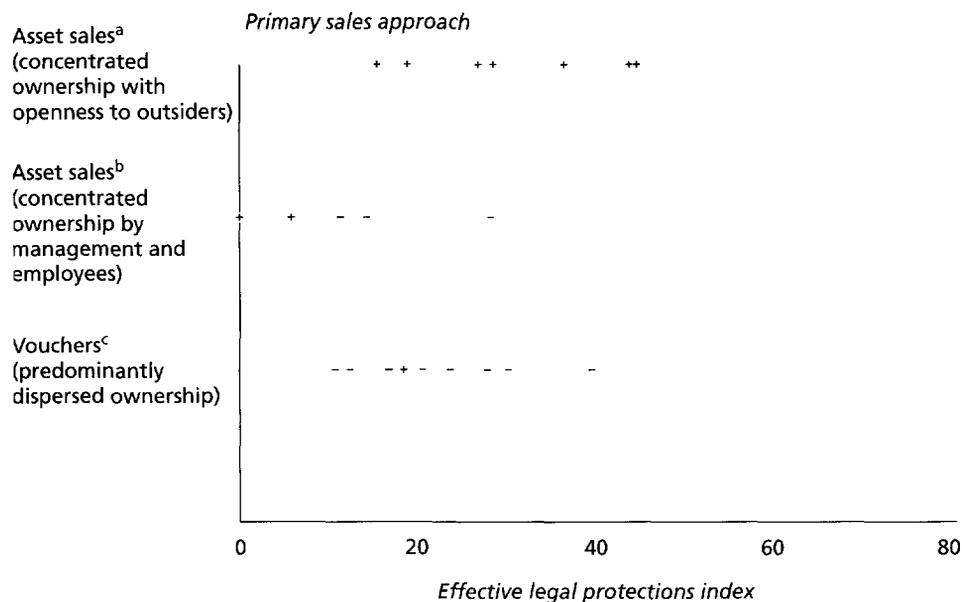
The transition economies followed neither established publicly traded firms nor privatized firms in the rest of the world. Here initial ownership structures were not related in any systematic way to the effectiveness of initial legal protections. In contrast with the privatization samples discussed above, many transition economies with weak effective legal protections used a privatization approach (vouchers) that did not facilitate concentrated ownership initially, and even those using asset sales largely eschewed attempts to sell firms to owners with an established reputation of promise fulfillment (figure 3). The presence of countries in the bottom left corner of figure 3 shows that they followed a path without empirical support and with little theoretical backing.

At the same time, several transition economies with weak effective legal protections at the time of privatization used asset sales open to outsiders. Estonia's privatization program, for example, closely followed that of Germany (see Nellis 1996). Vouchers were used, but in most instances minority stakes of only about 40 percent were targeted to voucher holders, while 60 percent was sold to strategic investors. Most strategic investors were foreign, with Swedes and Finns of Estonian descent playing an important role. The German experience is also illustrative. German privatization officials had significant freedom to choose the degree of ownership concentration, with many demands for a broad distribution of shares to eastern Germans. However, as Dyck (1997) describes, their approach—using asset sales rather than share issues, with openness to foreigners and a preference for firms with management capabilities and experience in the sector—resulted in eastern German firms being purchased by established western German firms. Eastern German firms thus inherited the corporate governance structure of established western German firms, a structure viewed internationally as effective in addressing governance problems.

Legal Protections and Financial Performance in Privatized Firms

Evidence from various studies is broadly consistent in its portrayal of the benefits associated with privatization. As Megginson and Netter (forthcoming) emphasize in a comprehensive survey of privatization, the financial impact of privatization is overwhelmingly positive. There is also evidence that privatization produces more effective corporate governance. Cragg and Dyck (1999a, b) present evidence that managers in privatized firms are far more likely to be rewarded or penalized on the basis of their financial performance than managers in state enterprises. Studies suggest that such governance reforms contribute to improved performance. The notable exception remains the transition economies, where performance has been

Figure 3. Effective Legal Protection and Dominant Approach to Ownership Concentration in Privatized Firms in Transition Economies



Note: + indicates that the country's GDP growth was higher than the mean for the region during 1991–98. – indicates that the country's GDP growth was lower than the mean for the region during 1991–98. The effective legal protection index is based on Pistor, Raiser, and Gelfer (2000) minority shareholder protection index in the year that marked the beginning of the country's privatization program. Regions include Central and Eastern Europe, the Baltics, and the Commonwealth of Independent States. a. Includes from left to right Eastern Germany, Kazakhstan, Slovak Republic, Estonia, Hungary, Poland, Latvia. b. Includes from left to right Slovenia, Belarus, Uzbekistan, Ukraine, Romania. c. Includes from left to right Azerbaijan, Russian Federation, Georgia, Kyrgyz Republic, Armenia, Moldova, Czech Republic, Lithuania, Bulgaria. Source: EBRD (1999) with author modifications. Growth data and regional classifications are from the International Monetary Fund as presented in Havrylyshyn and Wolf (1999).

disappointing and firms were more likely to adopt ownership structures not predicted by the extent of legal protections.

No systematic efforts have been made across countries that have pursued privatization to see whether the initial ownership structure at the time of privatization relative to the legal environment helps explain subsequent performance. But regional studies focusing on the transition economies find links between initial ownership structures and subsequent performance, measured by financial returns and the effectiveness of restructuring efforts.

An indication of a relationship between ownership structure and performance is provided by data on national level performance (this is just suggestive, many other factors determine national performance). The initial approach to privatization correlates with the country's subsequent growth experience. In figure 3 I have indicated with a plus sign countries that have per capita GDP growth rates from 1991 to 1998 that exceeded the mean for those in the same region. I have indicated with a nega-

tive sign countries that had lower per-capita GDP growth rates than countries in their region.¹¹ Those that held to international benchmarks and introduced concentrated ownership structures to reflect the weak legal environments have done better than average in their regions. Those countries that pursued voucher privatization, such as the Czech Republic, have done worse, sometimes spectacularly worse.

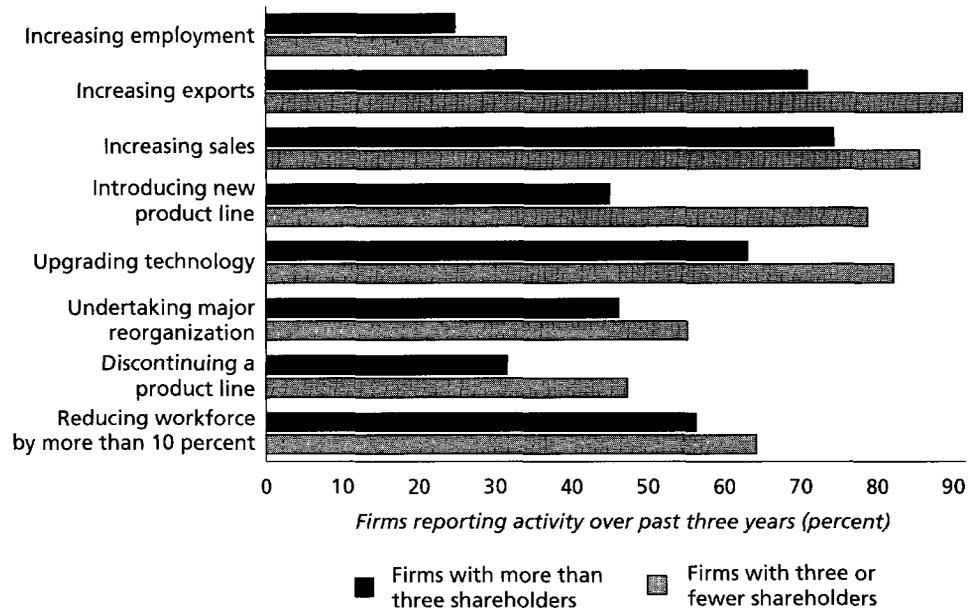
More convincing than this macroeconomic evidence are the results from firm-level studies. Djankov and Murrell (forthcoming) summarize evidence from 23 such studies in a meta-analysis, developing estimates of the impact of different ownership structures on performance relative to the alternative of continued state ownership. These studies, many of which control for country-specific factors and possible selection biases in the privatization method for particular firms, include Djankov (1999) for the Commonwealth of Independent States (CIS) and Frydman and others (1999) for the Czech Republic, Hungary, and Poland.

Their findings indicate that ownership structure has systematic and significant effects on performance. Consistent with the argument above and the evidence for established firms, in countries with weak legal protections dispersed ownership is the least effective corporate ownership structure. Compared with the most effective structure, foreign ownership, dispersed ownership has just a tenth the impact on performance. Outsiders are more effective than insiders in improving performance, with foreigners the most effective. Other outsiders with concentrated ownership, such as investment funds and blockholders, are also very effective.

A survey of 3,000 enterprises conducted by the World Bank and the EBRD provides similar findings (EBRD 1999). This study is less sophisticated in controlling for possible biases than the studies used in the Djankov and Murrell meta-analysis, but it has the advantage of a consistent data collection procedure. For all indicators of restructuring—from reducing the workforce to introducing new products and technology to increasing sales and employment—reform is much greater in firms with three or fewer shareholders than in those with more than three (figure 4). Beyond ownership concentration, the identity of the owners also matters for improved governance (figure 5). Firms sold to foreigners outperform state-owned firms and firms sold to domestic citizens on all indicators of restructuring. The stronger effects of ownership structure on performance in the institutionally weaker CIS countries in both figure 4 and in the Djankov and Murrell survey hints at the links between institutional development and the efficiency of different ownership structures.

Offsetting benefits to ownership structures that are inappropriate for the extent of legal protections are hard to find. Pistor, Raiser, and Gelfer (2000) present evidence that transition economies that pursued mass privatization started with stronger legal protections (2.55 on the anti-director rights index in 1992) than others (1.85) and have maintained the difference (3.61 in 1998, compared with 2.71). But the mass privatizers often introduced the improvements in protections after privatization rather than before, and the possible benefits of these legal protections appear to have been undermined by poorer market oversight. Pistor and others' index reveals that despite starting with stronger oversight in 1992, countries that had pursued mass privatization had weaker oversight by 1998.

Figure 4. Ownership Concentration and Restructuring Activity in Transition Economies



Source: EBRD (1999, chart 9.3), based on World Business Enterprise Survey of industrial firms (excluding new enterprises) with more than 200 employees.

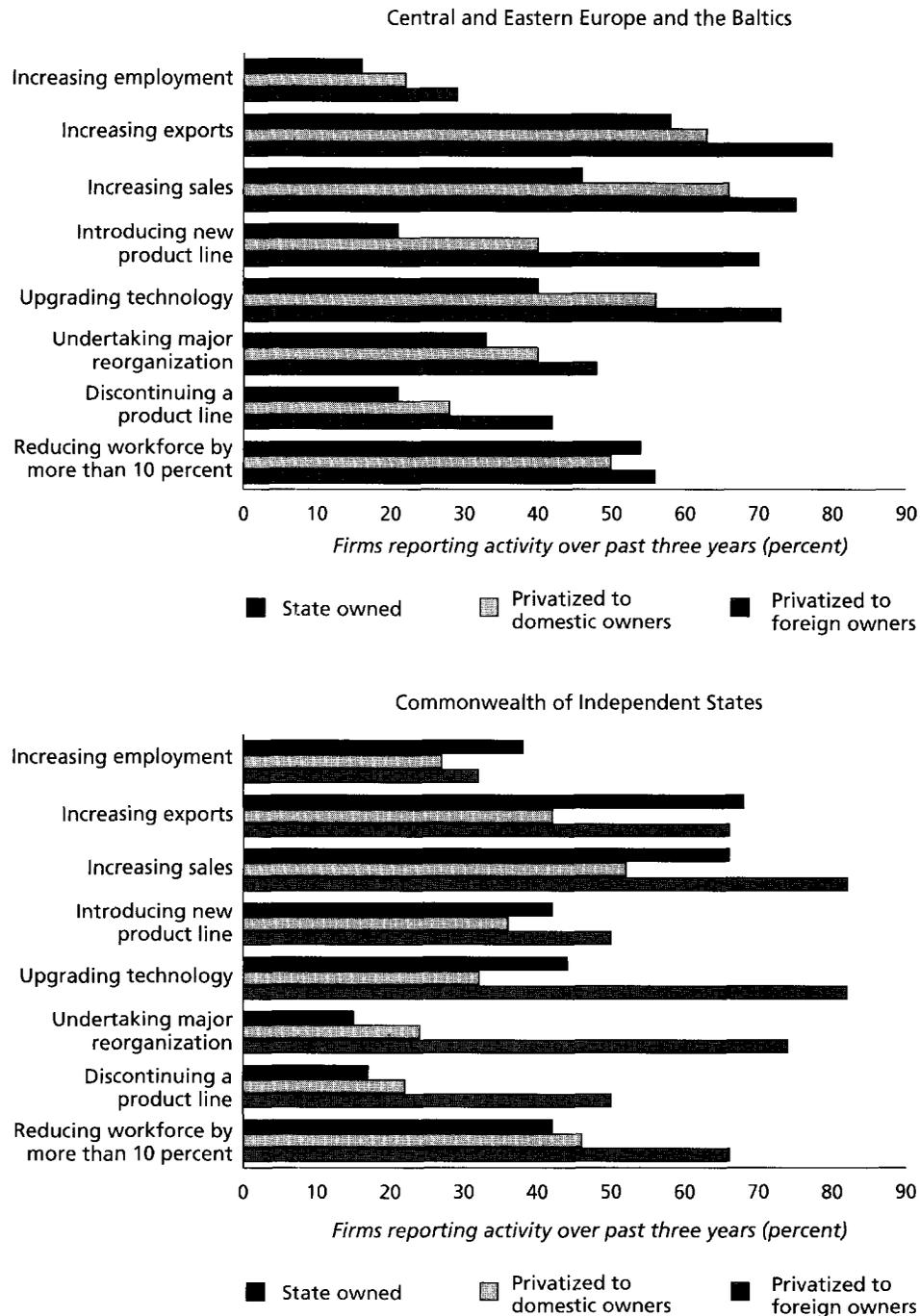
Policy Implications

What priorities do this functional framing and evidence suggest for decision makers interested in improving corporate governance systems? I present three broad agenda items, in declining order of importance.

Aligning Governance Policies with Existing Legal Protections

Governance policies relating to the ownership structure of privatized and established firms should be aligned with the existing state of legal protections. Institutions are amenable to change, but the institutional depth required to create effective legal protections develops only in the medium to long run. My reading of the evidence, particularly for privatized firms, is that a misalignment of ownership structure with the legal environment has large costs. In environments with weak legal protections, not only are anonymous, dispersed ownership structures unsustainable, but recognition of their lack of sustainability encourages socially wasteful activities. A possible rationale for such an approach—that it will stimulate an increase in effective legal protections—does not have much support.

Figure 5. Type of Ownership and Restructuring Activity in Transition Economies



Source: EBRD (1999, chart 9.2), based on World Business Enterprise Survey of industrial firms (excluding new enterprises) with more than 200 employees.

For established firms (such as firms affiliated with business groups), the privatization evidence suggests caution in adopting policies that transform established ownership patterns. Increasing the concentration of cash flow ownership and reducing the gap that may exist between cash flow ownership and voting rights (by limiting pyramiding, for example) are likely to be beneficial. But attempts to force dispersed structures are problematic. Although anonymous, dispersed shareholding has benefits relative to other ownership structures, those benefits do not outweigh the costs where legal protections are weak.

I do not disagree with scholars who see concentrated ownership structures as problematic. For example, Claessens, Djankov, and Lang (2000) suggest that:

The concentration of corporate control in the hands of a few families creates powerful incentives and abilities to “lobby” government agencies and public officials for preferential treatment...A concentrated control structure of the whole corporate structure could lead to the suppression of minority rights and hold back the institutional development of legal and regulatory channels to enforce those rights. (p. 109)

But there is no convincing evidence that forcing a replacement of such structures with more dispersed shareholding, while legal protections remain ineffective, will represent an improvement. In my reading of the evidence, the ownership structure is more a response to the institutional environment than the source of that environment.

Policy actions that produce marginal improvements in developed economies might have the opposite effects in economies with poor investor protections. Simply put, economists’ intuition based on the assumption that contracts will be upheld can be misleading in environments with weak legal protections. This suggests a need for caution in adopting common approaches to addressing governance concerns. Policymakers are advised to focus on fostering the functions of effective governance systems rather than replicating institutions.

Benchmarking Reforms of Governance Institutions

Reforms of governance institutions need comprehensive study and benchmarking to be effective. Institutional reforms have value, but because institutions are complementary, reforms have greater value when they occur together. So, the aim of corporate governance reform should not be to increase legal protections alone, but to increase effective legal protections that can provide the functions of governance. A dramatic improvement in laws will have little value if it is not enforced. Weaker legal reforms that can be enforced are more likely to limit abuses of authority by entrepreneurs.

The research described here has made significant progress in outlining tools for benchmarking. For legal protections, the technique of La Porta and others (1998) for identifying key legal protections for financiers is a good starting point, although, as the authors acknowledge, it is incomplete. Other legal protections not found in company law can substitute for the protections they discuss to lower the cost of

resolving competing claims on the wealth generated in the firm. Expanding the set of legal protections, as Pistor, Raiser, and Gelfer (2000) have done, is worth considering. Whether this set should include protections for other stakeholders in the firm remains an open question. Theoretically, labor, management, and suppliers, like financiers, condition their willingness to provide resources on their expectations about whether the explicit and implicit promises exchanged for these investments will be fulfilled. But whether legal protections for these investments should be included is an empirical question—one that remains to be investigated.

For the judiciary, summary measures of efficiency, as viewed by participants in the system or commercial agencies, provide a strong signal, perhaps the best available. The historical evidence from industrial economies and the correlation between common law tradition and some efficiency measures suggests that such features as constitutions, the distribution of powers among branches of government at the federal level, the degree of decentralization of federal authority, and the degree of devolution of power and taxing authority to regional and local levels may matter.

There is ample scope for better measures. Pistor and Wellons (1998) provide some alternatives. In their study of law and Asian economic development, they collected information on the time it takes to resolve cases, the willingness to take cases to court, and the fraction of cases in which the ruling went against the sovereign. Also useful would be indicators showing what structural variables at the discretion of policymakers contributed to judicial efficiency.

In most developing economies efforts to develop information intermediaries with well-aligned incentives are likely to rely on intermediaries headquartered in other countries, create conditions to facilitate the entry of foreign intermediaries, or transform existing closed intermediaries, such as banks, so that they produce information for public purchase. Chang, Khanna, and Palepu (2000) use data on the extent and accuracy of analyst activity across the world to benchmark such activity. The incentives of such intermediaries will be determined by the quality of regulation of financial intermediaries and by product market competition among intermediaries.

A few observations flow from the complementarity of these institutions and the need to consider their interaction. That the depth of complementary institutions differs across countries suggests that efforts to devise a checklist of legal protections that should be introduced everywhere are misguided. While there may be a common set of legal protections to which countries can aspire, reforms of legal protections must take into account a country's existing institutions to be most effective. Where complementary institutions are inefficient, little should be expected from changes in legal protections. Scarce policy resources would be better directed to developing these complementary institutions or to improving governance through other approaches.

Creating Alternative Mechanisms to Facilitate Promise Fulfillment

For countries with weak legal protections, an alternative mechanism for facilitating promise fulfillment might be piggybacking on institutions in more developed markets. One approach that works by identifying domestic companies with well-functioning for-

eign networks is to have such companies cross-list their shares on foreign exchanges. Once identified with a foreign exchange, the firm is no longer anonymous. Foreign companies listing on U.S. stock exchanges, for example, are required to comply with international accounting standards and to increase their information disclosure. Equally important, the exchange presumably has the incentive to delist firms if they violate their commitments to the exchange—a costly outcome if firms value the option to raise future capital and have few other channels to international capital. By allying themselves with the exchange, firms thus strengthen beliefs about promise fulfillment.

Privatized firms have led the way in cross-listing, accounting for more than a third of the \$233 billion raised through ADR programs from 1990 to 1999 (IMF 2000). Reese and Weisbach (2000) provide evidence that more broadly cross-listings have increased most among firms based in countries with weak legal protections. Johnson and Shleifer (1999) emphasize the growth of the Neuer Markt in Germany, a creation of the Frankfurt Stock Exchange with higher listing requirements that has served as the vehicle for many new listings of German and neighboring country companies.

While the early evidence is promising, there are limitations to this approach. It is of little help for resolving competing claims on the wealth generated by the firm; disputes are still resolved in the firms' home country courts, which tend to be inefficient. Moreover, questions arise about the penalty associated with delisting. This penalty is significant only if those controlling the firm value continued access to finance through that exchange and believe that delisting by one foreign exchange will reduce the possibility of listing on others. The fact that there are multiple exchanges—and little coordination among them—reduces the effectiveness of delisting as a penalty.

A second way forward is to foster domestic networks, such as business associations, that could increase the provision of information about firms and impose penalties on firms that violate promises when courts are arbitrary or unlikely to side with investors. Business associations are likely to be effective particularly if they require a large bond to join and if continued membership is valuable. In these circumstances, just knowing that a firm belongs to an association would strengthen investors' belief that the firm will uphold its promises.

But this approach, too, has limits. There are legitimate questions about whether associations have sufficient incentives to impose penalties on their members for non-compliance. In addition, firms with no interest in complying have incentives to pretend to comply—by forming a competing business association, for example. Where such mimicry is possible, and the real characteristics of a business association are known only in the long run, the prospects for such associations will be reduced. Moreover, reputation mechanisms are extremely sensitive to changes in future growth prospects. An economic shock that puts firms at risk exposes them to possible unrestrained looting.

Conclusion

Policymakers are right to concern themselves with systems of corporate governance. The evidence suggests that when the functions of a corporate governance

system are not provided, or are provided at high cost, firms violate promises and resources for new investment projects dry up, slowing national growth and development.

I have suggested the importance of governance institutions emphasizing the centrality of legal protections, which shift power to investors in firms and lower the costs of resolving disputes over the wealth of corporations. But as the evidence I have presented shows, the effectiveness of legal protections depends on complementary institutions that provide incentives for enforcement and ensure the provision of information to investors. Thus reforms that focus only on legal rules have limits, and universal recommendations for governance reforms will inevitably be inadequate. Improving the effectiveness of legal protections requires changes to multiple institutions, with the emphasis depending on a country's institutional endowment.

I have also highlighted ownership structure, which becomes an important instrument for providing the functions of governance in countries with ineffective legal protections. Evidence for both established and privatized firms reveals systematic links between particular ownership structures and financial performance. Evidence also shows that insights on ownership structure from industrial economies do not easily translate to settings with ineffective protections.

Theory and practice suggest a great many possibilities for tying the "grabbing hands" of public and private actors in private sector firms. The place to start is to focus on the functions of corporate governance.

Notes

1. Caves (1989), for example, describes the modern analytical economic approach as one that "is treated by its practitioners as institution free, exposing the consequences of fundamental human motives and technological opportunities unclouded by any detritus of law, culture, language, custom, or history" (p. 1226).

2. This approach draws from the functional perspective to the financial system developed in Crane (1995).

3. My approach closely follows that of Williamson (1985) and particularly Zingales (1997). Zingales defines corporate governance as "the complex set of constraints that shape the ex post bargaining over the quasi-rents generated by the firm" (p. 499).

4. My focus on the security of the promises made to all investors in the firm is broader than the definition of Shleifer and Vishny (1997), for whom "corporate governance deals with the ways in which suppliers of finance to corporations assure themselves of getting a return on their investment." (p. 737).

5. I borrow this phrase from Greif (1997), who emphasizes the importance of this difference in his analysis of medieval trade.

6. In the spirit of recent work by Greif (1997), I describe institutions as social factors (viewed as exogenous to investors in firms) that affect beliefs about transactions and produce regularities in behavior yet are endogenous to societal decisionmaking. Not all institutions of society are institutions of corporate governance. This label is reserved for institutions that are central to the governance transaction of investment in corporations.

7. I thank Rafael La Porta for encouraging me to explore this example in more detail and my colleague Tom McCraw for helpful clarifying discussions.

8. Among other factors, the SEC required a 20-day cooling-off period to scrutinize documents before issuing securities. It also had the power to place a stop order, which simply suspended the issue. This was a simple but powerful tool, as it shattered investor confidence in the issue whether or not the SEC was overturned on appeal.

9. In the community responsibility system, for example, investors know that they can demand payment from any member of that community, so they have little incentive to examine the capabilities of the controller of the investment or the viability of the proposed investment project. Bad projects will be pursued, and the community will be punished only after the fact.

10. The choice of countries reflects two restrictions: a country had to have five publicly traded companies without significant government ownership, and it could not be a formerly socialist country.

11. Regions used are eastern Europe, Baltics and the Commonwealth of Independent States. Growth data from IMF presented in Havrylyshyn and Wolf (1999).

References

- Aghion, Phillippe, and Olivier Blanchard. 1996. "State Firms in the Transition: On Insider Privatization." *European Economic Review* 1 (1): 759–66.
- Aoki, Masahiko. 1984. *The Cooperative Game Theory of the Firm*. Oxford: Oxford University Press.
- Bebchuk, Lucian. 1999. "A Rent Protection Theory of Corporate Ownership and Control." NBER Working Paper 7203. National Bureau of Economic Research, Cambridge, Mass.
- Becht, Marco, and Ailsa Roell. 1999. "Blockholdings in Europe: An International Comparison." *European Economic Review* 43 (4–6): 1049–56.
- Berglöf, Erik, and Ernst Ludwig von Thadden. 1999. "The Changing Corporate Governance Paradigm: Implications for Transition and Developing Countries." In Boris Pleskovic and Joseph Stiglitz, eds., *Annual World Bank Conference on Development Economics 1999*. Washington, D.C.
- Berle, Adolf, and Gardiner Means. 1933. *The Modern Corporation and Private Property*. New York: Macmillan.
- Black, Bernard, Reinier Kraakman, and Anna Tarassova. 1999. "Russian Privatization and Corporate Governance: What Went Wrong?" *Stanford Law Review* 52 (July): 1731–1804.
- Bolton, Patrick, and Ernst Ludwig von Thadden. 1998. "Blocks, Liquidity, and Corporate Control." *Journal of Finance* 53 (1): 1–25.
- Caves, Richard. 1989. "International Differences in Industrial Organization." In R. Schmalensee and R. Willig, eds., *Handbook of Industrial Organization* 2. Amsterdam: North Holland.
- Chang, J., Tarun Khanna, and Krishna Palepu. 2000. "Analyst Activity around the World." Harvard Business School, Cambridge, Mass.
- Chhibber, Pradeep, and Sumit Majumdar. 1999. "Foreign Ownership and Profitability: Property Rights, Control, and the Performance of Firms in Indian Industry." *Journal of Law and Economics* 42 (1): 209–38.
- Claessens, Stijn, and Simeon Djankov. 1999. "Ownership Concentration and Corporate Performance in the Czech Republic." *Journal of Comparative Economics* 27 (3): 498–513.
- Claessens, Stijn, Simeon Djankov, and Larry H.P. Lang. 2000. "The Separation of Ownership and Control in East Asian Corporations." *Journal of Financial Economics* 58 (2): 498–513.
- Coffee, John C., Jr. 1999. "The Future as History: The Prospects for Global Convergence in Corporate Governance and Its Implications." Working Paper 144. Columbia University School of Law, New York.

- Cragg, Michael, and Alexander Dyck. 1999a. "Management Control and Privatization in the United Kingdom." *Rand Journal of Economics* 30 (3): 475–97.
- . 1999b. "Privatization and Management Incentives: Evidence from the United Kingdom." Harvard Business School, Cambridge, Mass.
- Crane, Dwight, ed. 1995. *The Global Financial System: A Functional Perspective*. Boston, Mass.: Harvard Business School Press.
- Demsetz, Harold, and Kenneth Lehn. 1985. "The Structure of Ownership: Causes and Consequences." *Journal of Political Economy* 93 (6): 1155–77.
- Djankov, Simeon. 1999. "Ownership Structure and Enterprise Restructuring in Six Newly Independent States." *Comparative Economic Systems* 41 (1): 75–95.
- Djankov, Simeon, and Peter Murrell. Forthcoming. "Enterprise Restructuring in Transition: A Quantitative Survey." *Journal of Economic Literature*.
- Dyck, I. J. Alexander. 1997. "Privatization in Eastern Germany: Management Selection and Economic Transition." *American Economic Review* 87 (4): 565–97.
- . 2001. "Privatization and Corporate Governance: Principles, Evidence and Future Challenges." *World Bank Research Observer* 16 (spring): 59–84.
- EBRD (European Bank for Reconstruction and Development). 1999. *Transition Report 1999*. London.
- Frydman, Roman, Cheryl Gray, Marek Hessel, and Andrzej Rapaczynski. 1999. "When Does Privatization Work? The Impact of Private Ownership on Corporate Performance in the Transition Economies." *Quarterly Journal of Economics* 114 (4): 1153–91.
- Greif, Avner. 1997. "Informal Contract Enforcement: Lessons from Medieval Trade." In *The New Palgrave Dictionary of Economics and the Law*. London: Macmillan.
- Hart, Oliver. 1995. *Firms, Contracts and Financial Structure*. Oxford: Oxford University Press.
- Havrylyshyn, Oleh, and Thomas Wolf. 1999. "Growth in Transition Countries 1991–1998: The Main Lessons." Paper presented at the conference A Decade of Transition, International Monetary Fund, 1–3 February, Washington, D.C.
- Hong, Harrison, Jeffrey Kubik, and Amit Solomon. 2000. "Security Analysts' Career Concerns and Herding of Earnings Forecasts." *Rand Journal of Economics* 31 (1): 121–44.
- Hoshi, Takeo, Anil Kashyap, and David Scharfstein. 1990. "The Role of Banks in Reducing the Costs of Financial Distress in Japan." *Journal of Financial Economics* 27 (1): 67–88.
- IMF (International Monetary Fund). 2000. *International Capital Markets: Development, Prospects and Key Policy Issues*. Washington, D.C.
- Jensen, M. C. 1991. "Eclipse of the Public Corporation." In D. A. Oesterle, ed., *The Law of Mergers, Acquisitions, and Reorganizations*. St. Paul, Minn.: West.
- Johnson, Simon, and Andrei Shleifer. 1999. "Coase v. the Coasians: The Regulation and Development of Securities Markets in Poland and the Czech Republic." Harvard Institute of Economic Research Working Paper 1885. Cambridge, Mass.
- Johnson, Simon, John McMillan, and Christopher Woodruff. 1999. "Contract Enforcement in Transition." Stanford Graduate School of Business Working Paper. Stanford, Calif.
- Johnson, Simon, Peter Boone, Alasdair Breach, and Eric Friedman. Forthcoming. "Corporate Governance in the Asian Financial Crisis, 1997–1998." *Journal of Financial Economics* 58 (2): 141–86.
- Johnson, Simon, Rafael La Porta, Florencio López-de-Silanes, and Andrei Shleifer. 2000. "Tunnelling." *American Economic Review* 90 (2): 22–27.
- Kaplan, Steven. 1994. "Top Executives, Turnover and Firm Performance in Germany." *Journal of Law, Economics and Organization* 10 (1): 84–105.

- Kaplan, Steven, and Bernadette Minton. 1994. "Appointments of Outsiders to Japanese Boards: Determinants and Implications for Managers." *Journal of Financial Economics* 36 (2): 225–58.
- Khanna, Tarun. 2000. "Business Groups and Social Welfare in Emerging Markets: Existing Evidence and Unanswered Questions." *European Economic Review* 44: 748–61.
- La Porta, Rafael, Florencio López-de-Silanes, and Andrei Shleifer. 1999. "Corporate Ownership around the World." *Journal of Finance* 54 (2): 471–517.
- La Porta, Rafael, Florencio López-de-Silanes, Andrei Shleifer, and Robert W. Vishny. 1997. "Legal Determinants of External Finance." *Journal of Finance* 53 (1): 1131–50.
- . 1998. "Law and Finance." *Journal of Political Economy* 106 (6): 1113–50.
- . 1999a. "Investor Protection and Corporate Valuation." Harvard Institute of Economic Research Working Paper 1882. Cambridge, Mass.
- . 1999b. "Investor Protection: Origins, Consequences, Reform." Harvard University, Department of Economics, Cambridge, Mass.
- . 2000. "Agency Problems and Dividend Policies around the World." *Journal of Finance*. 55 (1): 1–33.
- López-de-Silanes, Florencio. 1997. "Determinants of Privatization Prices." *Quarterly Journal of Economics* 112 (4): 965–1025.
- McCraw, Thomas K. 1982. "With Consent of the Governed: SEC's Formative Years." *Journal of Policy Analysis and Management* 1 (3): 346–70.
- McMillan, John, and Christopher Woodruff. 1998. "Interfirm Relationships and Informal Credit in Vietnam." University of California at San Diego, Department of Economics.
- Meggison, William L., and Jeffrey M. Netter. Forthcoming. "From State to Market: A Survey of Empirical Studies on Privatization." *Journal of Economic Literature*.
- Meggison, William L., Robert C. Nash, Jeffrey M. Netter, and Annette B. Poulsen. Forthcoming. "The Choice of Privatization: An Empirical Analysis." *Journal of Financial Economics*.
- Merton, Robert C., and Zvi Bodie. 1995. "A Conceptual Framework for Analyzing the Financial Environment." In Crane and others, eds., *The Global Financial System: A Functional Perspective*. Boston: Harvard Business School Press.
- Morck, Randall, Andrei Shleifer, and Robert Vishny. 1986. "Management Ownership and Market Valuation: An Empirical Analysis." *Journal of Financial Economics* 20: 293–315.
- Nellis, John. 1996. "Finding Real Owners: Lessons from Estonia's Privatization Program." World Bank, Private Sector Development, Washington, D.C.
- North, Douglass, and Barry Weingast. 1989. "The Evolution of Institutions Governing Public Choice in 17th-Century England." *Journal of Economic History* 49 (1): 803–32.
- Olegario, Rowena. 1999. *Whom to Trust: The Dynamics of Business Lending in Nineteenth-Century America*. Ann Arbor: University of Michigan Business School.
- Pistor, Katharina, and Phillip A. Wellons. 1998. *The Role of Law and Legal Institutions in Asian Economic Development*. Oxford: Oxford University Press.
- Pistor, Katharina, Martin Raiser, and Stanislaw Gelfer. 2000. "Law and Finance in Transition Economies." *Economics of Transition* 8 (2): 325–68.
- Rajan, Raghuram, and Luigi Zingales. 1998. "Financial Development and Growth." *American Economic Review* 88 (3): 559–86.
- . 2000. "The Great Reversals: The Politics of Financial Development in the 20th Century." University of Chicago, Graduate School of Business Administration, Chicago, Ill.
- Reese, William, and Michael Weisbach. 2000. "Protection of Minority Shareholder Interests, Cross-Listings in the United States, and Subsequent Equity Offerings." Tulane University, Department of Economics, New Orleans, La.

- Shleifer, Andrei, and Robert W. Vishny. 1997. "A Survey of Corporate Governance." *Journal of Finance* 52 (2): 737–83.
- Weinstein, David, and Yishay Yafeh. 1998. "On the Costs of a Bank-Centered Financial System: Evidence from the Main Bank Relations in Japan." *Journal of Finance* 53 (2): 635–72.
- Williamson, Oliver. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.
- Wurgler, Jeffrey. 2000. "Financial Markets and the Allocation of Capital." *Journal of Financial Economics* 58(1): 187–214.
- Zingales, Luigi. 1997. "Corporate Governance." In *The Palgrave Dictionary of Economics and the Law*. New York: The Stockton Press.

Corporate Governance and Restructuring: Lessons from Transition Economies

G rard Roland

This article paints a synthetic picture of corporate governance in transition economies with an emphasis on its implications for efficiency, taking into account the heterogeneity of managers' skills, the diversity of firms' restructuring tasks and financial situations, and political constraints. The focus is on the efficiency effects of different privatization policies, with special emphasis on the broad dynamic effects of privatization. Different privatization policies have different effects on the distribution of economic power, with potentially far-reaching consequences for state capture, law enforcement, tax collection, stock market development, and private sector development. Moreover, legal reform is not an exogenous but an endogenous process influenced by the vested interests created by a country's initial privatization policy.

Corporate governance has recently become a hot topic in Western economies—one that is widely debated both in the business world and in academic research (for a survey, see Maher and Anderson 1999). Corporate governance is even more important in transition economies, where the initial situation was state governance and where approaches to corporate governance will likely affect overall economic performance.

Evolving Concepts of Corporate Governance in Transition

Achieving adequate corporate governance is no easy task in transition economies. Because almost the entire economy is involved, solutions must be matched to the varying conditions of firms. Simple recipes and slogans may attract a lot of attention,

G rard Roland is professor of economics at the University of California, Berkeley, and the European Centre for Advanced Research in Economics and Statistics at the Universit  Libre de Bruxelles. He is a visiting professor at the Center for Economic Research and Graduate Education–Economics Institute in Prague, program director for Transition Economics at the Centre for Economic Policy Research and Statistics, and a research fellow at the William Davidson Institute at the University of Michigan Business School. The author is grateful to Erik Bergl f for many conversations on the subject of this article. Part of the research for this paper benefited from an ACE grant from the European Commission.

Annual World Bank Conference on Development Economics 2000

 2001 The International Bank for Reconstruction and Development / THE WORLD BANK

but they are likely to give partial and sometimes misleading solutions. Early on in the transition, the slogan heard most often was “get the state out of the economy.” This sounded reasonable.

More recently, the focus has been less on the inefficiency of state-owned enterprises and more on financial scandals—the various forms of asset stripping and looting in privatized firms. This focus is especially visible in the Czech Republic and Russia, where mass privatization programs transferred huge parts of the economy into private hands at dazzling speeds. Given the various forms of “tunneling” through which majority shareholders in mass-privatized firms were able to strip assets and cheat minority shareholders, the more recent slogan has been “introduce the common law system.” In a celebrated series of papers, La Porta and others (1998a, b, 1999, 2000) have made an empirical connection between the type of legal system and the degree of shareholder protection. A key finding is that the common law systems in the United Kingdom and the United States protect minority shareholders more than the civil law systems in France, Germany, and other advanced economies of Western Europe.

This shift between such different slogans offers food for thought to those who have been uncritically espousing these slogans one after the other. When the emphasis was on getting the state out of the economy, fixing the law was a low-priority task and many legal loopholes were left open. In the Czech Republic, for example, former Prime Minister Václav Klaus put little emphasis on legal reform and financial sector regulation early in the transition. This was not necessarily because he believed that this approach would be good for tunneling but because he was wary of state intervention and thought that market development required as little state development as possible—a widely shared view in the early 1990s.

The current emphasis on corporate governance in transition is related to dissatisfaction with the simple dichotomy between state ownership and private ownership. The slogan “get the state out of the economy” partly relied on the idea that any form of privatization would always be much better than state ownership and that market mechanisms would lead to a more efficient distribution of assets among private owners.

But privatization has led to many disappointments. There is now extensive evidence on asset stripping by insiders to the detriment of minority shareholders (see Black, Kraakman, and Tarassova 2000; Coffee 1999). Other inefficiencies also persist, including soft budget constraints (Fox and Heller 1999; Schaffer 1998). And a large empirical literature has shown that in many privatized firms—especially firms privatized to insiders—there has been little restructuring or investment.

Moreover, the behavior of state enterprises has proven less apocalyptic than was initially predicted by many analysts, who foresaw massive asset stripping unless firms were privatized at great speed.¹ Early evidence on state enterprises actually showed cases of restructuring despite an absence of privatization (Belka, Krajewski, and Pinto 1993; Estrin, Schaffer, and Singh 1993).

The differences in enterprise behavior, whether privatized or not, can be related to the variety of approaches to corporate governance and to the variety of initial conditions facing enterprises. Is privatization to insiders always efficient? Under what conditions is it inefficient? Are inefficiencies of the same type, or do different

inefficiencies depend on specific conditions? How are these inefficiencies best corrected? These questions and others need to be answered to correctly assess situations and formulate policy recommendations. That is why this article draws a synthetic picture of corporate governance in transition economies with an emphasis on its implications for efficiency. The attention paid to scandals, frauds, and the like, especially in the legal literature, should not hide other inefficiencies that are perhaps less salient from a legal perspective but no less costly in terms of economic efficiency.

The main idea in this article is that the efficiency of corporate governance in transition economies is directly related to the privatization policies chosen and to the distribution of economic power and the economic environment generated by those policies. More broadly, the vested interests created by the initial distribution of economic power following various privatization policies are likely to have far-reaching consequences for state capture, law enforcement, tax collection, private sector development, and stock market development. In other words, privatization policies affect many variables that are important for overall economic performance.

Corporate Governance and Efficiency

Here corporate governance is defined as the control rights that influence the decisions of enterprise managers and ensure that outside funds can be raised to implement those decisions. (For closely related definitions, see Zingales 1998 and Williamson 1985.)

Problems with State Governance

Corporate governance in transition economies initially involved state governance—with many associated inefficiencies. A growing literature has tried to understand these inefficiencies.

A first strand of papers emphasized government intervention that imposes on managers political objectives such as the excess creation of jobs that offer political benefits but generate economic inefficiencies (Shleifer and Vishny 1994). State intervention also undermines commitment to incentive schemes. In contracts between private parties, governments (and courts) usually act as a third party to ensure contract enforcement and commitment. But in the relations between governments and enterprises, there is no such third party—so the question of government commitment to enterprises is important. Even if a government cares about efficiency, its inability to commit may lead it to intervene in enterprises, creating inefficiencies.

In the literature this has been called the ratchet effect (Laffont and Tirole 1988; Roland and Sekkat 2000). In its most general form the ratchet effect means that government uses new information generated by firm behavior to renege on its past commitments and “ratchet up” its demands on enterprises. Well-known examples are the ratcheting up of plan targets for enterprises that fulfill the plan too easily or the cutting of budgets for enterprises or divisions that do not spend their entire budgets.

The perverse incentives generated by the ratchet effect are well understood: agents exhibit slack performance or overspend their budgets.

A second strand of papers has emphasized inefficiencies associated with state governance through efforts by enterprises to secure soft budgets (Kornai 1980; Dewatripont and Maskin 1995) and to seek rents.

The important question is, how does privatization solve these problems? The question is not trivial. Indeed, all these problems can exist with private firms. For example, the state can impose on private firms objectives that deviate from efficiency—as when inefficient regulations are enforced. Similarly, government can renege on its commitment not only to state enterprises but also to private enterprises. An obvious example comes from taxation: governments rarely commit to tax rates. And soft budget constraints and rent seeking are not unique to state enterprises. Private firms also lobby to get subsidies and to seek rents.

Economic theory has been trying to understand why these problems are less prevalent in private firms than in state-owned firms. Shleifer and Vishny (1994) show that state intervention is more costly in private firms than in state-owned firms because in private firms, deviations from efficiency imply profit losses for the owners. The managers of state enterprises, by contrast, depend on the state for their livelihood anyway and do not face such high opportunity costs from accepting state intervention. Thus private owners must be paid more to deviate from profit objectives.

Similarly, government's commitment problem can be made less severe through markets and free mobility. Roland and Sekkat (2000) show how the existence of a managerial labor market due to the presence of a private sector can break government's monopsony power over managers and so make managers less dependent on government. The literature also shows that entry and competition harden budget constraints because they reduce the opportunity cost to government of not bailing out firms in terms of job losses and other costs of hard budget constraints (Dewatripont, Maskin, and Roland 2000; Maskin and Xu 1999; Roland 2000). Finally, the profit incentives of private firms can be stronger than the rent-seeking incentives (and so discourage the latter) when the profit opportunities are big enough.

These are just a few examples. Much research still needs to be done to better understand the deeper reasons for inefficiencies in state enterprises. But the current literature goes beyond the simple dichotomy between private and state ownership, seeking to understand the exact mechanisms through which private firms can be more efficient.

Problems with Dispersed Share Ownership

It has long been known, at least since Grossman and Hart (1980), that dispersed share ownership is associated with a free-rider problem in monitoring. Dispersed outside ownership has been an inherent feature of voucher privatization in transition economies. In the Czech Republic all citizens were given vouchers to buy shares

of enterprises, and no mechanism was put in place to encourage the emergence of strong block-holders. Poland's mass privatization program tried to avoid the dispersed share ownership of the Czech plan by having large investment funds hold shares of enterprises. But that did not solve the problem: dispersed share ownership simply emerged in the ownership of the investment funds, because citizens were given vouchers to buy shares in the funds instead of in firms.

If firms have dispersed ownership, even the best legal environment will not eliminate the free-rider problem and so improve management. Thus efforts to perfect the law can be misguided when dispersed ownership is widespread—as during mass privatization. The law can still be useful in protecting minority shareholders from abuse by management. But that will only happen if the law can be enforced, a fundamental issue discussed below.

Complete concentration of assets among shareholders also has its costs, because excessive monitoring may stifle initiative by management (Burkart, Gromb, and Pannunzi 1997). Some mix of concentrated and dispersed ownership trades off optimally the costs of free riding with the costs of excessive monitoring.

Problems with Lack of Minority Shareholder Protection

Minority shareholder protection has become an important issue in transition economies—especially in the Czech Republic and Russia, which relied on mass privatization. As noted, the term tunneling was coined to denote the ways in which managers (as majority shareholders) can strip assets to the detriment of minority shareholders.

Mass privatization tends to enhance the benefits of asset stripping. Managers or owners who have received an asset nearly for free can reap immense gains from asset stripping, even when doing so is not optimal from an economic perspective. Two questions arise. Why would managers want to engage in inefficient asset stripping once they own part of a firm? And why would mass privatization lead to more asset stripping than would sales, since the price paid for assets is a sunk cost?

The basic answer lies in the selection effects that result from different forms of privatization. Under competitive sales (like auctions and tender sales), assets are allocated to the owner with the greatest willingness to pay.² The selected owner is thus the most likely to maximize the value of the firm and not to engage in inefficient asset stripping. Such owners will never want to engage in actions that reduce the value of the assets below their purchase price. In other words, they will not be interested in purchasing firms whose value they could only reduce.

Under mass privatization, incumbent managers get control and are not necessarily likely to maximize the value of the assets. Whereas an outside owner might prefer to restructure the firm, an incumbent owner may prefer to engage in asset stripping. Such owners have stronger incentives to strip assets if the general economic environment is uncertain. Incumbents who become owners through mass privatization may then prefer to reap the sure gains from stripping assets rather than the uncertain benefits from investing in restructuring. Such uncertainty can even be

endogenous and lead to multiple equilibria. Managers who believe that there may be renationalization or political instability may engage in asset stripping, reinforcing political instability.

Detailed analyses of tunneling techniques have sometimes led to the conclusion that better laws could reduce tunneling. Indeed, every tunneling technique is associated with a loophole in the law—so it can be argued that closing those loopholes will reduce tunneling. But while good laws are necessary safeguards, they are not sufficient. Even if the law is improved, how can one be sure that it will be enforced, that insiders will not bribe regulators, and so on? For example, Russia has fairly good laws. But law enforcement is rather weak (see EBRD 1999; Berglöf and von Thadden 2000; and Black, Kraakman, and Tarassova 2000).

Access to Outside Funds

Control rights and their distribution change the incentives within firms, but access to outside funds is equally crucial to achieve efficiency. Usually, those with control rights over a firm bring outside funds or have the capacity to do so. This is not the case with mass privatization, whether to outsiders or insiders: the transfer of ownership is not linked to the financing of restructuring. As a result assets may be transferred into private hands that do not have easy access to outside funds.

Access to outside funds depends on the state and liquidity of the financial system and on whether corporate governance arrangements convince investors to put money in the firm. From that perspective, outside investors may ask for control rights in exchange for funds. This should be seen as a positive development because such a transfer of control rights may correct the initial allocation of assets achieved through privatization, reducing the possible insider bias. But situations differ, as discussed below. Not all firms need outside funds, even when outside control enhances efficiency. Nor will they necessarily relinquish control rights in exchange for outside funds.

Corporate Governance Systems

Let us now take a broader look at corporate governance systems. As noted, recent innovative work by La Porta and others (1998a, b, 1999, 2000) emphasizes the effects on corporate governance of different legal systems. The authors show that countries with civil law systems tend to have less protection of minority shareholders, leading to a higher concentration of shares and to less liquidity in the stock market.

This finding can be linked to previous research by Mayer (1990). He emphasized the difference between the Anglo-Saxon market-oriented financial system, based on the stock market and arm's-length banking, and the German or Japanese bank-oriented system, with less development of the stock market but with strong banks having long-term relationships with their clients. This research shows that corporate governance may be related to the structure of both the legal system and the financial system.

Scholars have recently tried to understand the institutional complementarities in various economic systems. For example, Pagano and Volpin (1999) find a negative correlation between shareholder protection and worker protection: low shareholder protection may be associated with high worker protection, and vice versa. The authors argue that a political coalition may form between managers and workers. Because managers benefit from low shareholder protection, they may offer higher worker protection in exchange for workers' support. This complementarity may explain why corporate governance institutions may not necessarily converge. This idea will be important for future research.

There may be not only institutional complementarities but also institutional substitutes. Berglöf and von Thadden (2000) note that in systems with less-developed stock markets, stronger competition and the presence of small investors may lead to the formation of business groups that substitute for the stock market. Such business groups are thus an alternative way of generating funds. But why is there a lack of institutional market development in the first place? What is the role of history and of newcomers and latecomers in the development of corporate governance systems?³

When thinking about corporate governance systems, one must ask how they emerge and how they evolve. Transition is an interesting opportunity for understanding this evolution because the initial conditions are visible. I return to this issue later in this article. The hypothesis is that different privatization policies have played an important role in explaining the differences in corporate governance in transition economies with similar initial economic and political conditions.

Determinants of Enterprise Restructuring in Transition Economies

There are two aspects of restructuring in transition economies: a sectoral aspect and an enterprise-level aspect. From a sectoral point of view, transition must lead to the shrinking of some sectors (steel, heavy industry) and the expansion of others (services). This implies that the entry of new firms is likely to play an important role, on par with the restructuring of existing enterprises. And given the importance of sectoral reallocation, investment based on retained earnings is not sufficient to reallocate resources across sectors—pointing to the importance of access to outside funds.

Restructuring at the enterprise level, as explained in Roland (1997), depends on four factors: the degree of product competition, the skills of incumbent managers, the need for external finance, and the degree of firm independence from government.

Degree of Product Competition

Carlin, Van Reenen, and Wolfe (1995) show that firms with a dominant position tend not to restructure. From a theoretical perspective, this finding is not self-evident. The Schumpeterian view may lead to the opposite conclusion. Monopoly power may provide stronger incentives for restructuring because of the monopoly rents extracted by incumbent managers when they restructure. The absence of restructuring may then

be explained by the fact that managers in monopolistic firms are not residual claimants. Here privatization—even to incumbent managers—will help achieve more restructuring.

But it is also possible that monopolies do not restructure because the profits from doing so would be lower than the government subsidies monopolies can extract by maintaining loss-making operations (Segal 1998). In this case privatization does not foster restructuring. Indeed, incentives for rent seeking will remain intact and may even be reinforced if managers are residual claimants of the government subsidies.

To boost efficiency, then, firms should be demonopolized and entry should be encouraged. From a political economy point of view, the task is to overcome resistance to demonopolization. Good arguments have been made for demonopolizing before privatization: resistance to demonopolization can be more intense after monopolies have been privatized than before (Newbery 1991; Tirole 1991). This is a sequencing mistake. Russia made this mistake and is suffering the consequences. The task of demonopolization is somewhat orthogonal to privatization and corporate governance. I will not pursue it further, but its policy importance should not be underestimated.

Skills of Incumbent Managers

It is important to distinguish between two types of restructuring: defensive restructuring and strategic restructuring (Grosfeld and Roland 1997). Defensive restructuring refers to the shedding of redundant workers to cut losses—a painful process that requires determined management. But it can be achieved by incumbent managers, in principle without further help, and especially without outside funds. Strategic restructuring refers to the creation of a business strategy for enterprise expansion. Defensive and strategic restructuring thus require different skills. Strategic restructuring also requires outside funds when the business strategy cannot be financed internally.

To understand the heterogeneity in managerial skills, make a basic distinction between “good” and “bad” managers. Bad managers do not have the skills required for successful restructuring and are not competent enough to acquire them. Thus they are expected to lose their jobs and so lose from strategic restructuring. Good managers possess enough managerial or learning skills. Thus they are expected to gain from strategic restructuring. Of course, there is asymmetric information on managerial skills, and different privatization policies may screen good managers from bad (see below).

Because bad managers are assumed to lose from strategic restructuring, they have no incentive to engage in defensive restructuring. Early privatization may even encourage bad managers to strip the assets of state enterprises (Aghion, Blanchard, and Burgess 1994). To boost efficiency, such managers must be replaced. The political economy constraint that must be overcome is resistance to layoffs within firms.

Good managers are expected to gain from strategic restructuring, but they may face different costs from defensive restructuring. These costs may be low enough for

some managers to perform defensive restructuring without being given formal incentives. Other managers may need such incentives. For efficiency, such managers must be given incentives for defensive restructuring.

Due to asymmetric information, there may be resistance to defensive restructuring both from bad managers and from good managers who face high costs from it. How can the good managers be encouraged to engage in defensive restructuring? Either by imposing negative incentives (such as the credible threat of closing the firm if it does not restructure) or by giving managers equity in the privatized firm. Both instruments screen good managers from bad.

Need for External Finance

An additional distinction should be made between firms with no retained earnings and firms with retained earnings. As noted, firms without retained earnings in which incumbent managers are expected to gain from strategic restructuring will need to raise outside funds. By doing so, they may give up control rights to investors. This may enhance economic efficiency within the firm because the transfer of control rights can lead to better decisions.

Firms with retained earnings have no need to raise outside funds. But giving away control rights can increase efficiency in such firms. This would be especially likely when incumbent managers engage in empire building and make excessive investments in the firm instead of diversifying investments.

In both cases efficiency requires securing outside finance and control. The political economy constraint is overcoming resistance from empire builders in firms with retained earnings.

Degree of Firm Independence from Government

As noted, major problems with state governance include inefficient government intervention in firms (the ratchet effect) and rent seeking by firms (the soft budget constraint). Here efficiency requires finding mechanisms for government commitment. It is necessary to overcome both forms of inefficiency, but one must think carefully about how to achieve that goal. For example, weakening the government will reduce intervention in firms—but possibly at the expense of increasing soft budget constraints. One must go beyond the simple dichotomy between a weaker and a stronger government and think of specific mechanisms to establish government commitment.

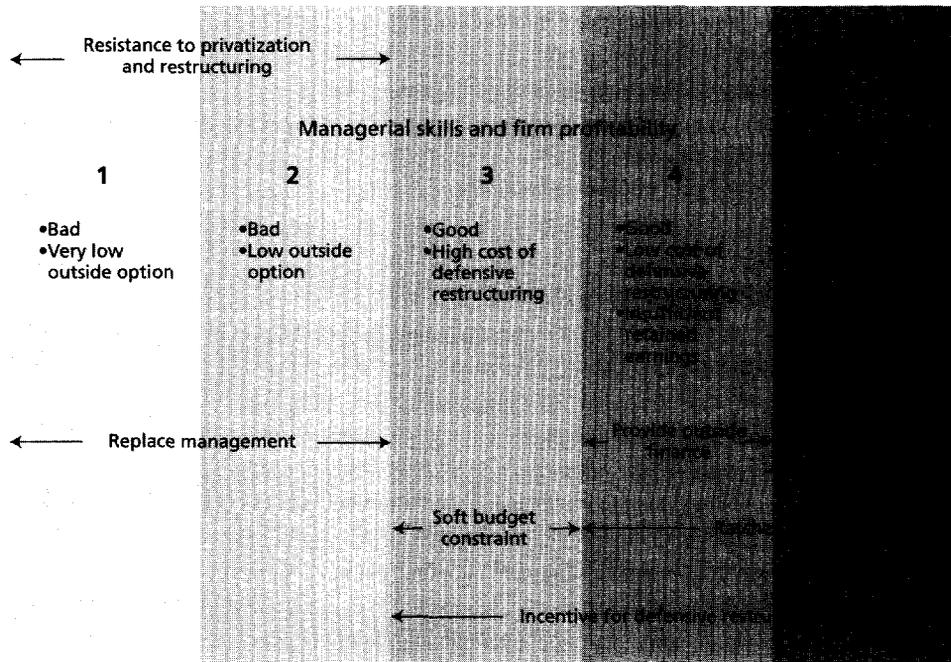
Summary

The preceding analysis holds for a given set of enterprise and sector characteristics. Controlling for those characteristics, it is reasonable to assume that there is a positive correlation between managerial skills and firm profitability. Figure 1 summarizes this in what I believe is the most comprehensive framework yet offered to explain the links between corporate governance and restructuring.

Thus there are various types of asymmetric information. There is the difference between bad and good managers: bad managers differ in their outside options, while good managers differ in their costs for defensive restructuring and in their need for outside funds. Overall, there are five types of managers. Type 1 are bad managers who are expected to lose from strategic restructuring and who have very few outside options. Type 2 managers are also bad but have more outside options. Thus these two types differ only in their outside options—a useful distinction. Indeed, because there is a political economy task of overcoming resistance to privatization and restructuring, it may be useful to develop tactics that prevent the joint resistance of the two groups (Dewatripont and Roland 1992, 1995). For efficiency, both types of managers should be replaced.

Type 3 managers are good but face a high cost for defensive restructuring. Types 4 and 5 are good and have a low cost for defensive restructuring. The difference is that type 4 managers do not have enough retained earnings to finance strategic restructuring, while type 5 managers do. Type 3, 4, and 5 managers need to be given incentives for defensive restructuring. But type 3 managers may need to be given stronger incentives. Under socialism, type 4 and 5 managers were more likely to suffer from the ratchet effect, while type 3 managers (as well as types 1 and 2) were more likely to be subject to soft budget constraints. For type 4 managers, efficiency requires securing outside finance. But for type 5 managers, it is also important to provide outsider control, though for a different reason: to avoid empire building. Type 5 managers also face the political economy constraint of overcoming resistance to outsider control.

Figure 1. Managerial Heterogeneity and Restructuring



Before analyzing the effects of different privatization policies on corporate governance and restructuring, one can use the framework of figure 1 to think about the different initial conditions in various countries. For example, some countries may have had more type 1 and 2 managers. Because proportions of manager types may differ across sectors and countries, policy priorities may not be the same everywhere.

Effects of Privatization Policies on Corporate Governance and Restructuring

Instead of discussing the optimal responses to the issues raised in the previous section, I will analyze the effects of different privatization policies on the five types of managers. There are four types of privatization policies:

- Fast giveaways (mass privatization) to dispersed outsiders (as happened in the Czech Republic).
- Fast giveaways to insiders (as in Russia).
- Top-down sales to outsiders (as in East Germany).
- Bottom-up gradual sales to outsiders (as in Hungary and Poland).

By top-down sales I mean those initiated primarily by the government. By bottom-up sales I mean those initiated through decentralized initiatives emanating from potential buyers or from managers seeking a buyer or major investor.

Fast Giveaways to Dispersed Outsiders

When there is a fast giveaway to dispersed outsiders, the free-rider problem in monitoring means that bad managers are unlikely to be replaced because dispersed shareholders are likely to be passive. In addition, type 4 managers are unlikely to get outside finance, and type 5 managers are unlikely to face outsider control. But the reasons are not exactly the same. Because the transfer of ownership and the provision of outside funds are decoupled in mass privatization, it may not be easy for a privatized firm to get outside funds for strategic restructuring. Small outside owners may lack the necessary capital, or dispersed ownership may discourage investors unless they gain control. On the other hand, type 5 managers will be left to engage in empire building because of the free-rider problem in monitoring.

But this type of privatization does give type 4 and 5 managers incentives for defensive restructuring because they will be residual claimants of the effort put into it. Similarly, this form of privatization eliminates the ratchet effect because the state loses control over managers. Type 3 managers will not necessarily have incentives for defensive restructuring because its cost is high relative to the uncertain benefits of future participation in a firm with outside finance. But the soft budget constraint may be partly solved for such firms. Indeed, mass privatization deprives the government of much of its wealth and so raises the cost of public funds, increasing the cost of bailing out firms.

Though this privatization policy may not achieve efficiency objectives, it is likely to ease political constraints. Bad managers will not be threatened by dispersed share-

owners and so will be less resistant to this form of privatization. The same reasoning applies to type 5 managers who do not need to fear outsider control.

Fast Giveaways to Insiders

A fast giveaway to insiders is likely to have effects quite similar to those of a fast giveaway to outsiders. The only difference is in terms of formal ownership. Under dispersed ownership, managers do not have ownership, but they have control. A fast giveaway to insiders gives managers real ownership on top of real control.

Top-Down Sales to Outsiders

Top-down sales to outsiders are likely to achieve all the efficiency objectives of privatization. New owners will replace type 1 and 2 managers and provide outside finance and outside control to type 4 and 5 managers. But political constraints will likely be a problem. Type 1, 2, and 5 managers will strongly resist this privatization policy, though for different reasons. Type 1 and 2 managers will resist because they want to avoid being ousted, while type 5 managers will resist in order to continue empire building.

It would be wrong to analyze privatization policies without taking into account political constraints. It is not by chance that top-down sales to outsiders were used only in East Germany, where political constraints were least important because the country was being merged with the bigger and powerful West Germany. In unified Germany, East Germans have had, until recently at least, little influence on political decisionmaking. Because of the differences in political constraints, it would be wrong to compare fast giveaways that achieve fewer efficiency objectives with top-down sales to outsiders.

Bottom-Up Gradual Sales to Outsiders

Bottom-up gradual sales to outsiders are easier to compare with mass privatization (see Bolton and Roland 1992). Under gradual sales, good firms with better managers tend to be privatized first. This is in line with political economy analysis (Roland 1994a, 1994b; Dewatripont and Roland 1997). In addition, better managers try to signal themselves to potential investors by engaging in defensive restructuring (Roland and Sekkat 2000). This is consistent with the observation that firms restructure to attract the attention of private investors (Estrin, Schaffer, and Singh 1993). Gradual sales are more likely to ease political constraints because type 1 and 2 managers will, at least during the early stages, not be the object of privatization. The efficiency objective of replacing them will thus not be obtained in the first stage. So, there will be no resistance to privatization and restructuring as long as those managers do not feel threatened and are not facing a direct privatization prospect. The problem of replacing bad managers will be tackled only later, after reformers have grown stronger.

But gradual sales are still likely to encounter resistance to outside control. The level of resistance will depend on initial conditions. Walsh and Whelan (1999) show

that Polish firms producing exports for the EU market are less subject to privatization than firms exporting to the former Council of Mutual Economic Assistance (CMEA) market. This seems to indicate that the first group of firms is more resistant to privatization and is exploiting the opportunities provided by the booming EU export market.

Gradual sales have definite efficiency advantages over mass privatization. Indeed, the need for outside financing and control is not decoupled from the transfer of ownership because the investors who buy firms have the means to restructure them. Otherwise they would not be buying them. And the fact that firms get privatized gradually allows for harder budget constraints—to the extent that negative incentives can be used against managers. Indeed, because type 1 and 2 managers are not threatened at an early stage, it is more credible for government to be hard on type 3 managers, in order to avoid soft budget constraints in those firms.

Summary

The effects on restructuring of the four privatization policies are summarized in table 1. The table includes a row on the policies' effects on asset stripping. As noted, firms privatized to dispersed outsiders or to insiders may encounter asset stripping. Gradual sales can prevent asset stripping in privatized firms but not necessarily in firms left under state control—especially those controlled by type 1 and 2 managers. A policy like East Germany's, with fast top-down sales to outsiders, prevents asset stripping more fully.

Although fast giveaways (mass privatization) take into account political constraints, they achieve only some of the efficiency objectives associated with privatization. This is also the case for gradual sales in their early stages. Though there are differences between the two methods—especially the better outside finance and control under gradual sales—the static analysis presented above indi-

Table 1. Outcomes of Privatization Policies

<i>Outcome</i>	<i>Fast giveaways (mass privatization) to dispersed outsiders</i>	<i>Fast giveaways (mass privatization) to insiders</i>	<i>Top-down sales to outsiders</i>	<i>Gradual bottom-up sales to outsiders</i>
Efficiency objectives				
Replace management	No	No	Yes	Yes
Engage in defensive restructuring	Partial	Partial	Yes	Gradual
Secure outside funds and control	No	No	Yes	Gradual
Harden budget constraints	Partial	Partial	Yes	Gradual
Prevent asset stripping	No	No	Yes	Partial
Political constraints satisfied				
Resistance to redundancies	Yes	Yes	No	Yes
Resistance of empire builders to outside control	Yes	Yes	No	No

cates that they should have similar outcomes in their early stages. Whether the differences become smaller or larger depends on a dynamic analysis of the effects of the various privatization policies.

Dynamic Effects of Different Initial Allocations of Economic Power

The most common argument about the dynamic effects of privatization policies used to be that the efficiency shortcomings of mass privatization would quickly disappear because of market forces. Firms privatized to incompetent insiders would soon be taken over by more efficient owners, and entrepreneurial insiders would soon find outside finance for strategic restructuring. But that has not happened. Insider entrenchment has occurred, preventing takeovers. Moreover, insiders who were expected to become powerful advocates for reform have become powerful interest groups, defending the rents they received through mass privatization.

Thus this section places more emphasis on the dynamic effects of the initial allocation of economic power generated by privatization policies. These dynamic effects shed light on the divergence between countries that used mass privatization as their main privatization policy and countries that relied on gradual sales. I am not claiming that different privatization policies have necessarily been the main determinant of different performance across transition economies. Rather, I want to highlight some of the implications of the distribution of economic power created by privatization policies.

As a starting point, note the big difference between the economic power of individuals under state and private ownership. Under state ownership in the socialist economy, the state was extremely powerful, holding nearly all political, military, and economic power. On their own, people like enterprise managers did not have much economic power: they were dependent on the planning system's chain of orders and commands. Under private ownership managers have far more economic power. Given that socialist enterprises were larger than capitalist firms, the person who gains economic power over a formerly socialist firm will have an enormous amount of power.

So, mass privatization creates a sudden and strong concentration of economic power among insider managers. In one move, managers are freed from the commands and constraints of a planned economy.

Gradual sales do not lead to such a sudden and strong concentration of power, because control over firms can be acquired only through sales and asset purchases (though that can occur through noncash sales such as leases or loans; see Bolton and Roland 1992). On the other hand, managers of state enterprises that have not been privatized do not have as much power as managers in insider-privatized firms—especially when the government reinstates its control rights over them, as in Poland (Belka, Krajewski, and Pinto 1993).

Given the powerful shift in economic power created by mass privatization, there is considerable scope for abuse of minority shareholders. Such abuse is especially likely if minority shareholders are voucher-holders who have no previous experience with stock markets and have fewer incentives to protect the value of their assets than normal small shareholders who have accumulated private savings.

This abuse of power will lead to low stock market confidence and low liquidity, shrinking the stock market. Between 1995 and 1999 the number of listings on the Prague (Czech Republic) stock exchange fell by more than 80 percent, from 1,716 to 301. Over the same period there was a 60 percent drop in the value of an index of the exchange's top 50 stocks (Coffee 1999).

When stock markets lack liquidity, insider-managed firms will try to form large business groups to pool financial resources and benefit from possible cross-subsidies. The formation of such groups can only lead to a further concentration of economic power.

The sudden shift of economic power to insider managers may also make it easier for them to capture the state. Insider managers can bribe politicians or use the threat of reducing economic activity and destroying jobs to extract subsidies or favorable legislation. Politicians tend to be more responsive to such influences when they emanate from those in control of large, visible firms. These influences will lead to more corruption within the state, weak tax enforcement (especially for large firms, a situation unique to transition economies; see Campos 1999), weak law enforcement, the development of mafia networks, and so on. Large insider interest groups may block legal reform that would reduce their power or undermine their interests. All these developments aggravate inflationary dangers in the economy.

One should not underestimate the economic power created by insider privatization. This power is greater than the private wealth of those insiders. For example, Russian oligarchs like Boris Berezovsky are said to have few direct stakes in firms but exercise strong control over them through control over their managers.

Over time this strong economic power is likely to lead to enormous wealth inequality (Alexeev 1999) because of asset stripping and capital evasion. Capital evasion will likely increase political instability given the high inequality in the distribution of wealth. This political instability will likely reinforce the short-term perspectives of managers and insider owners: they will prefer to strip assets rather than invest in the long-term future of the enterprises they control. At the same time, increased asset stripping will likely reinforce political instability—creating a vicious circle (Polishchuk 1999). Thus mass privatization can create a corporate governance system with no stock market liquidity, large business groups, weak law enforcement, and political instability.

In the Czech Republic these negative effects can be partly offset by prospects for joining the European Union. Those prospects may help generate a minimum of discipline in law enforcement and focus expectations in the right direction (Roland and Verdier 1999). But in Russia and other mass-privatizing countries with little hope of joining the European Union, the dynamic effects of mass privatization will likely have negative long-term effects.

Gradual sales, on the other hand, create a less sudden concentration of economic power. Because it takes money to acquire control, private capital must accumulate before more concentrated control over firms can be achieved. Much of this private accumulation is based on the new private sector and the small and medium-size enterprises that emerged at the start of the transition.

With gradual sales, the capital accumulated in small and medium-size enterprises will provide much of the new liquidity in the stock market. Thus these enterprises

become the small investors who have stakes in defending minority shareholder protection and who will lobby to change laws in that direction. Small and medium-size enterprises will also be the main constituency for further reform and for further shrinking of the state.

At the same time, under gradual sales state enterprises will continue to push for soft budget constraints. But these pressures can be eased through the growing strength of the private sector, which will increasingly oppose such practices as it grows in size and strength. Thus there may be a virtuous process leading to increased stock market liquidity, better minority shareholder protection, more law enforcement, real transparency in securities markets, further private capital accumulation, and stronger constituencies in favor of further reforms.

Empirical Evidence on Restructuring

There is an extensive empirical literature on the effects of the privatization form on enterprise performance. Before assessing that literature, it is worth noting that early studies on enterprise behavior found that the main difference was between existing firms (state-owned and privatized) and new enterprises (see Konings, Lehmann, and Schaffer 1996, Konings 1997, and Bilsen and Konings 1997). Controlling for factors such as size and capital intensity, new enterprises performed better on measures such as productivity growth. But such exercises are not easy to conduct because of a potential sample selection bias. The new enterprises generally survived early competition, which tends to overstate their performance relative to existing enterprises.

Another problem is the endogeneity that occurs when analyzing the effect of ownership form on enterprise performance by regressing performance on ownership. Causality may run the other way, with performance determining ownership—as when stronger enterprises are privatized first or privatized primarily to insiders who have superior information about them. In the Czech Republic the more profitable firms were privatized first to create support and political goodwill, in line with the political economy analysis mentioned above (Gupta, Ham, and Svejnar 1999).

Barberis and others (1996) provide an interesting empirical analysis of retail shops in Russia. Their study aims to provide evidence on the relative importance of the two channels mentioned earlier through which privatization increases efficiency: better matching of managers and assets, and better incentives. The authors find that restructuring occurred far less often when a shop was privatized to managers than when there were new owners who differed from incumbent managers. These findings suggest that matching managerial skills with assets is more important than improving incentives for incumbent managers. Incentives were found to have less effect than replacing management. The authors used an instrumental variable approach to control for endogeneity when using the method of privatization as an instrument. But while the study is interesting and well done, Russian shops provide limited evidence on the effects of privatization—especially the privatization of large firms.

Earle and Estrin (1997) surveyed 300 Russian enterprises after mass privatization was completed in mid-1994. The authors found first that the state remained a dom-

inant owner—holding more than 40 percent of shares—in 38 percent of the firms. More than 70 percent of the firms with a dominant private owner were privatized to insiders. The general finding from the authors' ordinary least squares regressions is that private ownership has a small positive impact on enterprise performance. The impact is big in cases of ownership by insiders and investment funds but insignificant for other forms of outside ownership.

The authors then correct for endogeneity (better firms being selected for privatization) using the privatization method as their instrument, as in Barberis and others (1996). The main difference obtained in the instrumental variable approach was that outside ownership has a significant positive effect on enterprise performance—except dispersed ownership, which has a negative effect. Thus the difference between the instrumental variable approach and the ordinary least squares approach suggests that better firms were privatized to managers and weaker ones to outside owners. But the performance measures relate to the period immediately after privatization, and further studies on Russia are needed—especially given subsequent events such as the generalization of barter and big increases in payment arrears.

Frydman and others (1999a) use panel data for 209 medium-size manufacturing firms (100–1,500 employees) in the Czech Republic, Hungary, and Poland in the fall of 1994. About 25 percent of the firms were owned by insiders, 25 percent by foreigners, 20 percent by investment funds, and 15 percent predominantly by the state. Using a fixed effect estimation, the authors find that outsider privatization had a positive effect on revenue and productivity growth and that insider privatization helped reduce layoffs. Among outside owners, foreigners contributed significantly to revenue and employment growth, while domestic financial firms and large minority state ownership increased revenue and productivity growth. The latter result is rather astonishing and suggests state passivity in those firms. Ownership by domestic outsiders as individuals had no effect on performance. Privatization had no effect on cost cutting, which suggests that privatization is more important for strategic restructuring than for defensive restructuring and that defensive restructuring can occur before privatization.

Using the same data, Frydman and others (1999b) show that major product restructuring has an important positive effect on revenue for outsider-owned firms but not for insider- and state-owned firms. This effect was strong even in outsider-privatized firms where management was not replaced. More broadly, the two papers by Frydman and others confirm the importance of the distinction made by Grosfeld and Roland (1997) between defensive and strategic restructuring. Defensive restructuring is as effective in privatized enterprises as in state enterprises. But strategic restructuring (as measured by revenue growth) is affected by ownership. Outsider-privatized firms are much better at strategic restructuring than state enterprises or insider-privatized firms.

Grosfeld and Nivet (1997) examine the impact of defensive and strategic restructuring on a large panel of Polish firms through 1994. The authors find that privatized firms invested a lot more and grew faster, confirming the effect of privatization on strategic restructuring. The authors do not distinguish between types of privatization. But in their sample, privatization occurred through sales to strategic investors.

Marcincin and van Wijnbergen (1997) find that in the Czech Republic, firms privatized entirely through vouchers tended to be of lower initial quality than firms privatized through sales or partial sales. Controlling for the selection effect, the authors find that privatization to outsiders had a positive but small impact on performance and that firms privatized entirely through vouchers exhibited weaker performance.

Weiss and Nikitin (1998) also analyze privatization in the Czech Republic. They find that concentrated outside ownership improves enterprise performance only under outside ownership other than investment funds. This result is even stronger when changes in performance are regressed against changes in ownership. Doing so eliminates any initial selection effect but is a noisy measure of control because dominant owners may reduce their ownership stakes after having secured control. The authors partly attribute the weak performance of firms owned by investment funds to perverse corporate governance arrangements and to asset stripping.

Prasnikar and Svejnar (1998) use data from Slovenia to test whether asset stripping was occurring. They test whether managers of state enterprises who also owned private firms underinvested in the state enterprises—and find that was not the case. They find that outside ownership had a positive effect on investment but no effect on wage setting. This suggests that privatization has not reduced the bargaining power of workers.

The European Bank for Reconstruction and Development's Transition Report 1999 provides preliminary evidence on the broader rent-seeking environment of privatization by examining the link between state governance, the degree of state capture, and the results of privatization. Businesses were surveyed to get an indirect measure of state capture by powerful interests. Firms were asked whether the sale to private interests of parliamentary votes and presidential decrees had affected their business. More than 40 percent of the firms in Azerbaijan, Moldova, Russia, and Ukraine cited a significant impact, but less than 10 percent in Slovenia and Uzbekistan did so. Thus there are "high capture" and "low capture" countries.

The influence on government appears to be concentrated in a small number of firms: in general less than 5 percent of the firms surveyed reported having a significant influence on policy. There is a negative correlation between the degree of capture and a general measure of governance (the effectiveness, as perceived by firms, of state governance in terms of regulation, taxation, inflation, policy stability, physical infrastructure, and law and order). Moreover, privatization's effect on governance differs according to the degree of capture. In low-capture states, progress on large-scale privatization is associated with better governance—while in high-capture states, progress on privatization is associated with worse governance. The second group is made up of Azerbaijan, Bulgaria, Croatia, Georgia, Moldova, Kyrgyz Republic, Romania, Russia, and Ukraine.

The evidence from transition economies tends to confirm prior analyses of privatization, indicating that privatization can enhance enterprise performance if corporate governance is sound. It is not true that any form of privatization is always better than state ownership. In particular, insider privatization in Central Europe and privatization to investment funds in the Czech Republic show disappointing results. And dispersed outside ownership can even hurt performance.

The empirical literature tends to show the importance of using privatization to better match managerial skills with assets. But many of these analyses were conducted just after privatization was completed. Empirical analysis over an extended period will improve our understanding of the policies followed. Comparative analyses across countries, like those of Frydman and others (1999a, b), are useful for assessing the overall process. Further research is needed to better assess the impact of the privatization policies followed in various countries, controlling for country-specific characteristics. Such cross-country analysis is also needed to understand the impact of the legal, financial, and political environment. Finally, research is needed to understand channels for hardening budget constraints through privatization. For example, has insider privatization helped harden budget constraints? If so, how?

Conclusion

A country's privatization policy appears to play a crucial role in determining restructuring outcomes within firms. Privatization policies also affect economic performance and the evolution of institutions and of reform. Differences between privatization policies matter because they affect the distribution of economic power among economic agents.

Mass privatization creates a sudden and strong concentration of economic power in the hands of insider managers. This sudden transfer of power creates opportunities for asset stripping and abuse of power to the detriment of minority shareholders. These abuses lead to low stock market liquidity and low confidence. When financial markets are weak, large business groups tend to form—reinforcing the concentration of economic power and leading to captured governments and potential political instability.

Gradual sales yield a lower initial concentration of economic power and rely on the entry of new entrepreneurs, learning of entrepreneurial skills, and accumulation of private capital. Small and medium-size enterprises and middle-class citizens become constituencies in favor of further reform and stable democracy. Their growing economic and political power makes them advocate for legal safeguards against asset stripping. Capitalism evolves naturally instead of by jumpstart.

Notes

1. The implicit assumption behind such reasoning was that all managers were bad and had to be replaced. Seen from today's perspective, this is clearly a sweeping generalization.
2. In transition economies those with the greatest willingness to pay are not necessarily those with the greatest ability to pay. Thus cash sales may allocate assets only to those with the greatest ability to pay. Solutions to overcome that problem have been put forward, particularly the concept of noncash sales (Bolton and Roland 1992) such as payment by installments, various leasing formulas, sales on credit, and gradual sales of stocks. Such noncash bids have been used extensively in Hungary and Poland; for a formal analysis, see Bolton, Pivetta, and Roland (1997).
3. This question was first raised by Gerschenkron (1962).

References

- Aghion, Philippe, Olivier Blanchard, and Robin Burgess. 1994. "The Behavior of State Firms in Eastern Europe Pre-privatization." *European Economic Review* 38: 1327–49.
- Alexeev, Michael. 1999. "Privatization and the Distribution of Wealth in Russia." *Economics of Transition* 7: 449–65.
- Barberis, Nicholas, Maxim Boycko, Andrei Shleifer, and Nadejda Tsukanova. 1996. "How Does Privatization Work? Evidence from Russian Shops." *Journal of Political Economy* 104 (4): 764–90.
- Belka, Marek, Stefan Krajewski, and Brian Pinto. 1993. "Transforming State Enterprises in Poland." *Brookings Papers on Economic Activity* 1. Washington, D.C.: The Brookings Institution.
- Berglöf, Erik, and Ernst-Ludwig von Thadden. 2000. "The Changing Corporate Governance Paradigm: Implications for Developing and Transition Economies." In Boris Pleskovic and Joseph E. Stiglitz, eds., *Annual Bank Conference on Development Economics 1999*. Washington, D.C.: World Bank.
- Bilsen, Valentijn, and Jozef Konings. 1997. "Job Creation, Job Destruction and Growth of Newly Established, Privatized and State-Owned Enterprises in Transition Economies: Survey Evidence from Bulgaria, Hungary and Romania." WDI Working Paper 106. University of Michigan Business School, William Davidson Institute, Ann Arbor.
- Black, Bernard, Reinier Kraakman, and Anna Tarassova. 2000. "Russian Privatization and Corporate Governance: What Went Wrong?" *Stanford Law Review* 52: 1731–1808.
- Bolton, Patrick, and Gérard Roland. 1992. "Privatization in Central and Eastern Europe." *Economic Policy* 15: 276–309.
- Bolton, Patrick, Frédéric Pivetta, and Gérard Roland. 1997. "Optimal Sale of Assets: The Role of Non-cash Bids." European Centre for Advanced Research in Economics and Statistics, Brussels.
- Burkart, Mike, Denis Gromb, and Fausto Pannunzi. 1997. "Large Shareholders, Monitoring and the Value of the Firm." *Quarterly Journal of Economics* 112 (August): 693–728.
- Campos, Nauro. 1999. "Never around Noon: On the Nature and Causes of the Transition Shadow." CERGE Discussion Paper 1999–19. Center for Economic Research and Graduate Education–Economics Institute, Prague.
- Carlin, Wendy, John Van Reenen, and Thomas Wolfe. 1995. "Enterprise Restructuring in the Transition: An Analytical Survey of the Case Study Evidence from Central and Eastern Europe." *Economics of Transition* 3 (4): 427–58.
- Coffee, John. 1999. "Privatization and Corporate Governance: The Lessons from Securities Market Failure." Columbia Law School, New York.
- Dewatripont, Mathias, and Eric Maskin. 1995. "Credit and Efficiency in Centralized and Decentralized Economies." *Review of Economic Studies* 62: 541–55.
- Dewatripont, Mathias, and Gérard Roland. 1992. "Economic Reform and Dynamic Political Constraints." *Review of Economic Studies* 59: 703–30.
- . 1995. "The Design of Reform Packages under Uncertainty." *American Economic Review* 85: 1207–23.
- . 1997. "Transition as a Process of Large Scale Institutional Change." In David Kreps and Kenneth Wallis, eds., *Advances in Economics and Econometrics: Theory and Applications*. Cambridge: Cambridge University Press.
- Dewatripont, Mathias, Eric Maskin, and Gérard Roland. 2000. "Soft Budget Constraints and Transition." In Eric Maskin and Andras Simonovits, eds., *Planning, Shortage and Transformation*. Cambridge, Mass: MIT Press.

- Earle, John S., and Saul Estrin. 1997. "After Voucher Privatization: The Structure of Corporate Ownership in Russian Manufacturing Industry." CEPR Discussion Paper 1736. Centre for Economic Policy Research, London.
- EBRD (European Bank for Reconstruction and Development). 1999. *Transition Report 1999: Ten Years of Transition*. London.
- Estrin, Saul, Mark Schaffer, and Inderjit J. Singh. 1993. "Enterprise Adjustment in Transition Economies: Czechoslovakia, Hungary and Poland." *Economic Journal on Eastern Europe and the Soviet Union* 3 (November): 21–51.
- Fox, Merritt B., and Michael A. Heller. 1999. "Lessons from Fiascos in Russian Corporate Governance." University of Michigan Law School, Ann Arbor.
- Frydman, Roman, Cheryl Gray, Marek Hessel, and Andrzej Rapaczynski. 1999a. "When Does Privatization Work? The Impact of Private Ownership on Corporate Performance in the Transition Economies." *Quarterly Journal of Economics* 114 (4): 1153–92
- . 1999b. "Why Ownership Matters? Politicization and Entrepreneurship in the Restructuring of Enterprises in Central Europe." New York University, Department of Economics, New York.
- Gerschenkron, Alexander. 1962. *Economic Backwardness in Historical Perspective*. Cambridge, Mass.: Harvard University Press.
- Grosfeld, Irena, and Jean-Fran ois Nivet. 1997. "Wage and Investment Behaviour in Transition: Evidence from a Polish Panel Data Set." CEPR Discussion Paper 1726. Centre for Economic Policy Research, London.
- Grosfeld, Irena, and G rard Roland. 1997. "Defensive and Strategic Restructuring in Central European Enterprises." *Journal of Transforming Economies and Societies* 3 (4): 21–46.
- Grossman, Sanford, and Oliver Hart. 1980. "Takeover Bids, the Free-rider Problem and the Theory of Corporation." *Bell Journal of Economics* 11: 42–64.
- Gupta, Nandini, John Ham, and Jan Svejnar. 1999. "Priorities and Sequencing in Privatization: Theory and Evidence from the Czech Republic." University of Pittsburgh, Department of Economics, Pittsburgh.
- Konings, Jozef. 1997. "Firm Growth and Ownership in Transition Economies." *Economics Letters* 55 (3): 413–18.
- Konings, Jozef, Hartmuth Lehmann, and Mark Schaffer. 1996. "Job Creation and Job Destruction in a Transition Economy: Ownership, Firm Size and Gross Job Flows in Polish Manufacturing 1988–91." *Labour Economics* 3 (3): 299–317.
- Kornai, J nos. 1980. *Economics of Shortage*. Amsterdam: North-Holland.
- Laffont, Jean-Jacques, and Jean Tirole. 1988. "The Dynamics of Incentive Contracts." *Econometrica* 51: 1153–75.
- La Porta, Rafa el, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny. 1998a. "Law and Finance." *Journal of Political Economy* 106: 1113–55.
- . 1998b. "Legal Determinants of External Finance." *Journal of Finance* 52: 1131–50.
- . 1999. "Corporate Ownership around the World." *Journal of Finance* 54: 471–517.
- . 2000. "Agency Problems and Dividend Policies around the World." *Journal of Finance* 55 (6): 1–33.
- Maher, Maria, and Terry Anderson. 1999. "Corporate Governance: Effects on Firm Performance and Economic Growth." Paper presented at the conference on Convergence and Diversity in Corporate Governance Regimes and Capital Markets, 4–5 November, Tilburg University, Tilburg, the Netherlands.
- Marcincin, Anton, and Sveder van Wijnbergen. 1997. "The Impact of Czech Privatization Methods on Enterprise Performance Incorporating Initial Selection-Bias Correction." *Economics of Transition* 5 (2): 289–304.

- Maskin, Eric, and Chenggang Xu. 1999. "Soft Budget Constraint Theories: From Centralization to the Market." Paper presented at the Fifth Nobel Symposium in Economics: The Economics of Transition, 10–12 September, Stockholm.
- Mayer, Colin. 1990. "Financial Systems, Corporate Finance, and Economic Development." In R. G. Hubbard, ed., *Asymmetric Information, Corporate Finance and Investment*. Chicago: University of Chicago Press.
- Newbery, David. 1991. "Reform in Hungary: Sequencing and Privatization." *European Economic Review* 35: 571–80.
- Pagano, Marco, and Paolo Volpin. 1999. "The Political Economy of Corporate Governance." Harvard University, Department of Economics, Cambridge, Mass.
- Polishchuk, Leonid. 1999. "Distribution of Assets and Credibility of Property Rights." University of Maryland and New Economic School, Moscow.
- Prasnikar, Janez, and Jan Svejnar. 1998. "Investment and Wages during the Transition: Evidence from Slovene Firms." WDI Working Paper 184. University of Michigan Business School, William Davidson Institute, Ann Arbor.
- Roland, Gérard. 1994a. "On the Speed and Sequencing of Privatization and Restructuring." *Economic Journal* 104: 1158–68.
- . 1994b. "The Role of Political Constraints in Transition Strategies." *Economics of Transition* 2 (1): 27–41.
- . 1997. "Economic Efficiency and Political Constraints in Privatization and Restructuring." Paper presented at the European Bank for Reconstruction and Development conference on Private Sector Development, November, London.
- . 2000. *Transition and Economics: Politics, Markets and Firms*. Cambridge, Mass.: MIT Press.
- Roland, Gérard, and Khalid Sekkat. 2000. "Managerial Career Concerns, Privatization and Restructuring in Transition Economies." *European Economic Review* 44 (10): 1857–72.
- Roland, Gérard, and Thierry Verdier. 1999. "Law Enforcement and Transition." European Centre for Advanced Research in Economics and Statistics, Brussels; Université Libre de Bruxelles; and DELTA, Paris.
- Schaffer, Mark. 1998. "Do Firms in Transition Economies Have Soft Budget Constraints? A Reconsideration of Concepts and Evidence." *Journal of Comparative Economics* 36 (1): 80–103.
- Segal, Ilya. 1998. "Monopoly and Soft Budget Constraints." *Rand Journal of Economics* 29 (3): 596–609.
- Shleifer, Andrei, and Robert Vishny. 1994. "Politicians and Firms." *Quarterly Journal of Economics* 109: 995–1025.
- Tirole, Jean. 1991. "Privatization in Eastern Europe: Incentives and the Economics of Transition." *NBER Macroeconomics Annual*. Cambridge, Mass.: National Bureau for Economic Research.
- Walsh, Paul, and Ciaara Whelan. 1999. "Firm Performance and the Political Economy of Corporate Governance: Survey Evidence for Bulgaria, Hungary, Slovakia and Slovenia." LICOS Discussion Paper. Catholic University of Leuven, Centre for Transition Economies, Belgium.
- Weiss, Andrew, and Georgi Nikitin. 1998. "Performance of Czech Companies by Ownership Structures." Boston University, Department of Economics, Boston.
- Williamson, Oliver. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.
- Zingales, Luigi. 1998. "Corporate Governance." In Peter Newman, ed., *The New Palgrave Dictionary of Economics and the Law*. Gordonsville, Va.: Palgrave/St. Martin's Press.

Comment on “Corporate Governance and Restructuring: Lessons from Transition Economies,” by Gérard Roland

John S. Earle

The dilemmas, policies, and outcomes of privatization in transition economies have fueled a growth industry in the analysis of corporate governance systems. The reasons for the interest in ownership and governance in transition are clear. Under socialism, dominant state ownership was complemented by tight bureaucratic control over large industrial firms—the “commanding heights” of the economy (Kornai 1992). After communism collapsed, questions immediately arose about what forms of ownership and control should follow and how the new system should be arranged.

The discussions of these issues have displayed a remarkable evolution in the 10-odd years since transition began. In the early transition years, the “Washington consensus” favoring rapid liberalization and privatization was dominant, while more recently a “new consensus” (Washington or not) has apparently emerged that advocates gradualism and cautious privatization.¹ Both viewpoints have had substantial effects on policymaking in the transition economies and on the advice provided by international organizations. But neither has ever been universally accepted, nor has the debate between them been informed by an amount of scientific analysis that would justify firm conclusions.

Limits of our Knowledge

What are the limits to our conceptual understanding and empirical knowledge of corporate governance and restructuring in the transition economies? To start, although theory has developed rapidly and in some exciting directions, it is still incapable of offering precise empirical predictions. Existing theory suggests few specific hypotheses, and it provides little guidance for empirical testing and quantification of the magnitude of effects.

Consider the problem of estimating the impact of privatization on firm restructuring and performance, one of the main topics in this literature and in the article by

John S. Earle is associate professor at the Stockholm School of Economics.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

Gérard Roland. Leaving aside how privatization, restructuring, and performance are measured—difficult enough questions about which theory has little to say—theory does not suggest a set of variables that should be controlled for in the analysis, nor does it imply a particular functional form for the relationship.

Such regressions, for instance as summarized in the review article of Djankov and Murrell (2000), routinely include control variables representing the initial quality of the firm and its retained earnings, an important issue developed in Roland's article. But the relationship between initial conditions and restructuring is complex. It includes the positive effect due to the availability of resources for restructuring and the negative one due to entrenchment, both of which are cited by Roland. But it also includes a negative effect due to reduced competitive pressure in firms with cash cushions and a positive effect due to avoiding the difficult restructuring associated with downsizing a firm. Thus the net effect is ambiguous, and no structural, behavioral model has accounted for these separate aspects of the relationship.

Empirical knowledge of the relationship between privatization and restructuring is limited by data and by the small number of studies that have been performed. One problem is the frequent use of data from early in the process. For example, Earle and Estrin (1997, 1998) analyze survey data from Russia collected in July 1994, just after the conclusion of voucher privatization. Many Russian firms were actually privatized earlier (in 1991–93), but it would certainly be desirable to collect more recent information to assess the impact of privatization. Until very recently, however, the data used in those analyses, collected through a World Bank research project, were the best available. To take another example, Frydman and others (1999) analyze data from 207 firms in the Czech Republic, Hungary, and Poland, measuring performance through 1993. Not only is the sample small, but again the data are far too dated to permit a confident evaluation of the effects of privatization.

Other problems involve sampling and questionnaire design. The Business Environment and Enterprise Performance Survey, a project by the World Bank and European Bank for Reconstruction and Development, gathered data on a large group of enterprises in 20 transition economies. But it employed quota rather than probability sampling, and it failed to collect information on previous ownership and other potential determinants of performance. As a result, analyses of the ownership-performance relationship using this database (appearing, for instance, in EBRD 1999) use contemporaneous ownership and past performance, calling into question the assumption maintained throughout the work that ownership may be taken as exogenous. Finally, a general problem in many studies is that data collection does not receive the careful attention to questionnaire design, interviewer training, and data entry and cleaning essential to ensuring reliable information.

Beyond the problems with data timing, sample sizes, sampling techniques, and data collection, a more fundamental issue in empirical work on the effects of privatization is establishing the appropriate counterfactual. Much of the "new consensus" contains strong assertions about the effects of privatization in various countries. But beyond the lack of empirical evidence to support most of those claims (which are not necessarily the fault of the authors, but rather a consequence of the state of

research in the area), a major problem is that we cannot observe what would have happened without privatization or with a different privatization policy.

Establishing the appropriate counterfactual is important both for estimation strategy (how to control for selection or simultaneity bias in the observed relationship) and for policy evaluation. Policies are sometimes infeasible for political reasons, and it makes little sense to compare an implemented policy with a Panglossian world where first-best policies can always be implemented (Shleifer and Treisman 2000). This is most obvious in Roland's discussion of privatization in East Germany, where all the "efficiency objectives of privatization were likely to be met." But it also holds in other circumstances.

Problems of Mass Privatization

A good example is the most forthright policy message in the "new consensus," the contention that mass privatization was a failure because it increased asset stripping and reduced restructuring. Before considering the alternatives to mass privatization, let us examine the elements of this proposition. In Roland's variant, mass privatization is defined as giveaways to dispersed outsiders and asserted to have the same effects as giveaways to insiders. But the definition of asset stripping is unclear. "Reducing the value of the assets" is not very persuasive as rational owners will always try to maximize the returns from their assets, so this does not differentiate asset stripping from restructuring. Moreover, there is an important difference between stripping assets and diverting cash flow, since the latter may maintain asset values while finding other uses for the returns to the assets.

The argument appears to assume that cash flow stripping and restructuring are substitutes in managerial effort, although this assumption is not supported by any empirical evidence. Indeed, the two actions may be positively correlated, as active owners try to improve the cash flow of their firms while doing their utmost to hide it from the tax authorities. A possible explanation is that "good" managers (in Roland's terminology) may be good not only at strategic restructuring but also at enriching themselves: thus the skills required for the two tasks are correlated.

Why is mass privatization alleged to increase asset or cash flow stripping? According to Black, Kraakman, and Tarassova (1999), granting control rights through the Russian privatization program increased the scope for theft. But this analysis neglects the problem that control rights were already *de facto* held by managers, and that privatization involved a transfer of legal cash flow rights. Roland argues that mass privatization tends to enhance the net benefit of asset stripping because managers or owners have received the asset nearly for free, thus they can reap immense gains from stripping it even when doing so is not optimal from an economic perspective. But the purchase price is a sunk cost from the new owner's perspective: because it is contracted before the owner has control rights it should not affect later behavior. Even an owner who has paid for an asset—or particularly, a partial share in an asset—may be able to gain more by reducing its value, say, by diverting cash flow (to avoid sharing rents with other shareholders or the state).

The usual selection arguments associated with well-organized auctions (that is, the new owner has the highest valuation) may also not apply. Such auctions may be difficult to organize in the poor informational and institutional environments of transition economies, and in fact very few countries have carried them out; more common for large and medium-size enterprises were tenders, where multiple criteria were employed in addition to price, and where corruption was rife. Furthermore, the usual efficiency proposition associated with ownership resulting from an auction may fail if ability to pay does not reflect valuation, and auctions may be politically infeasible where ability to pay is socially unacceptable as a criterion. This leads us to the question of counterfactuals.

Privatization Alternatives

What then is the realistic counterfactual to mass privatization? The new consensus does not make clear whether the claim of increased asset stripping under mass privatization is meant relative to continued state ownership or relative to some other form of privatization. Nor is any evidence cited on the amount of asset stripping actually observed under those alternatives. The discussion treats sales and giveaway privatizations as a dichotomy, but in reality privatization transactions involve a range of prices. Giveaways have seldom been completely free, and sales have rarely achieved what the assets might have fetched in a market setting—though we will never know because of valuation difficulties (see below).

A higher privatization price may reduce restructuring if it has a significant income effect on new owners with limited liquidity. Capital-constrained owners may be forced to liquidate assets to stay liquid. An example is the widespread unbundling of assets by Polish firms, partly to finance their privatization through leasing. Again though, it is unclear whether this should be classified as restructuring or asset stripping.

On the other hand, since the earliest discussions mass privatization has been criticized for not directly raising investment finance. But once again the counterfactual is not clear, for who had access to such finance? In the early 1990s most transition economies had little domestic savings (except Hungary). The possibilities that existed for financing privatization purchases were spread fairly evenly over the population except for some concentration in the nomenklatura (again, except in Hungary). Thus the only effective demand for large companies being privatized through sales came from the former communist elite and from foreigners.

Transferring loads of assets to the former elite would have resulted in incompetent owners and would have decreased the legitimacy of the process as much as any other program did (some of which did not differ much from such a direct transfer). And citizens of most countries feared foreigners because of nationalistic movements. For their part, foreigner investors were cautious given uncertainty about the reversibility of the process, the possibility of implicit or explicit renationalization, and the unknown usefulness of communist assets in a capitalist world.

Thus the general lack of demand for state assets in early transition was intertwined with two other problems: distributional conflicts and difficulty of valuation.

That the assets of state enterprises had been accumulated almost entirely during the socialist period—when costs, depreciation rates, and cash flow were almost meaningless—implied that standard accounting measures based on historical costs or projected cash flows were next to useless.

Cash auctions were feasible in principle, but they made it difficult to advance other social goals (ensuring investment, preserving employment, cleaning the environment). And foreigners and the former *nomenklatura* usually would win the auctions—giving rise to a highly skewed distribution favoring the two least popular groups in society. The importance of social goals led instead to sales through tender and direct negotiation in East Germany and Hungary. Given the nontransparency of these deals and the impossibility of reaching an objective valuation for each company, such privatizations were open to charges that the decisions reached on buyers and prices had been corrupt. Opposition parties regularly leveled such charges against incumbent governments, politicizing the process and making the bureaucrats in charge of sales still more timid.

As a result every country that pursued sales moved slowly, with the exception of East Germany. Private capital must be accumulated before more concentrated control over firms can be achieved, and that takes time. Except among the *nomenklatura*, no private savings will initially be available for large buyouts. If one assumes that private savings is initially equal to 6 percent of the corporate capital stock and grows by 10 percent a year, then it will take almost 30 years to privatize the stock.

The citizens of most transition economies did not want to wait that long. For one thing there were serious concerns about reverting to communism in the first few years after its collapse. In addition, citizens of transition economies wanted to deliver a functioning capitalist system if not to their children then at least to their grandchildren. And there was recognition that slow privatization would likely lead to steady depletion of the capital stock in state enterprises, given the breakdown in the state's ability to monitor them. Could the state have prevented this asset-stripping by continued socialist-style active administration of its enterprises while gradually privatizing and permitting new enterprises to enter, as suggested by Kornai (1990) and Murrell (1992)? To some extent this is what has been happening in China, although continued state control of the economy has also been associated with political control by the Communist Party (consistent with Kornai's 1992 analysis of the interdependence of the political, ownership, and coordination elements of classical socialism).

But the interesting counterfactual question is whether this setup was possible in Eastern Europe and the former Soviet Union. As with any counterfactual, we will never know the answer with certainty. But it is clear that no country managed to achieve this.

The possibility for the sales approach was greatest where the demand constraint was most relaxed due to greater private savings and availability of foreign capital, namely in Hungary. Moreover, Hungarian managers had more autonomy and experience in decisionmaking as a result of progressive liberalization in the 1970s and 1980s. Finally, Hungarian managers and consumers had acquired experience and

openness through small entrepreneurial activity permitted under the "New Economic Mechanism" and subsequent reforms, and from contact with foreigners, leading to foreign direct investment in the early 1990s that (even excluding privatization deals) dwarfed that of neighboring countries.

Thus Hungary had special advantages in pursuing sales (Earle, Frydman, and Rapaczynski 1993; Stark and Bruszt 1997). But these very advantages made sales more urgent by increasing the problem of asset stripping. Indeed, discussions of privatization in Hungary were dominated throughout the 1990s by the fear that the residual state sector was gradually leaching away under the influence of various forms of spontaneous privatization and asset diversion.

Related to the appropriate pace and method of privatization is the question of whether the state should restructure firms before releasing them to the private sector. One consideration concerns the situation when firms have market power, and privatization somehow permits that power to be exercised more freely. Roland takes a strong view, arguing that Russia made the mistake of privatizing before demonopolizing and now suffers the consequences. The argument appears to assume that Russian authorities were capable of carrying out an active, appropriately targeted, and uncorrupt competition policy. But were they? By all accounts, Russia's antimonopoly commissions were effective only at maintaining price controls over many firms, not at forcing splitups, and there have been charges of corruption in their choice of firms to regulate (Joskow, Schmalensee, and Tsukanova 1994).

Conclusion

So what lessons does transition teach about corporate governance and restructuring? First, notwithstanding all the controversy, most studies (including those cited in Roland's review of the empirical literature) do support the proposition that private ownership matters; despite data problems and the limitations of short time span, few studies find no effect of privatization. Second, it is critical to distinguish different methods of privatization, with their associated ownership structures and corporate governance outcomes. Third, privatization is not a panacea: the enormous restructuring problems in transforming enterprises will not be overcome easily or quickly, and ownership change is insufficient to keep the state from interfering with firms' decisions. Rather, privatization is only one of several important policies whose joint effects are complex (Brown and Earle 2001).

Fourth, privatization is not easily separated from new firm development. There have been few attempts to quantify the size of the new private sector, but much evidence points to antecedents in old state enterprises. While the relationship between privatization and new entry is frequently posed as a trade-off (see, for example, Murrell 1992), it might rather be complementary, as privatization policies set the basic conditions for startup entry. Finally, transition appears to show that ownership concentration has ambiguous effects—due not to excessive monitoring (Burkart, Gromb, and Pannunzi 1997) but to entrenched managers and suboptimal owners and to controlling owners that expropriate minority investors.

Most important, we still have a lot to learn, and we need a lot more evidence before we “rush to judgment,” as suggested by Stiglitz (2000). For an initiative that has attracted so much attention and controversy, privatization has benefited from relatively little systematic research. An evaluation requires better and more recent data than have generally been available, as well as careful attention to econometric problems that arise in estimating the relationship between enterprise ownership and performance. It also requires an analysis of ownership evolution and takeovers, which may render moot the initial postprivatization ownership structure. At a more basic level, we should distinguish between different privatization policies and components (for example, loans for shares and voucher privatization in Russia, or tunneling and the voucher scheme in the Czech Republic) and determine the appropriate counterfactual for the evaluation. While paying close attention to existing evidence, efforts to expand evidence remain essential—as does retaining scientific open-mindedness as new evidence emerges. The most important (and reliable) lessons from transition are yet to come.

Note

1. The early consensus in favor of rapid privatization is summarized by Lipton and Sachs (1990), Balcerowicz (1995), Blanchard and Layard (1992), Frydman and Rapaczynski (1991), and Boycko, Shleifer, and Vishny (1993). Critics of rapid privatization include Kornai (1990 and 2000); Murrell (1992); Stiglitz (1994, 2000); Black, Kraakman, and Tarassova (1999); and Spicer, McDermott, and Kogut (2000).

References

- Balcerowicz, Leszek. 1995. *Socialism, Capitalism, Transformation*. Budapest: Central European University Press.
- Black, Bernard, Reiner Kraakman, and Anna Tarassova. 1999. “Russian Privatization and Corporate Governance: What Went Wrong?” *Stanford Law Review* 52: 1731–1808.
- Blanchard, Olivier, and Richard Layard. 1992. “How to Privatize.” In Horst Siebert, ed., *The Transformation of Socialist Economies: Symposium 1991*. Tübingen: Mohr (Siebeck).
- Boycko, Maxim, Andrei Shleifer, and Robert W. Vishny. 1993. “Privatizing Russia.” *Brookings Papers on Economic Activity* 2: 139–81.
- Brown, J. David, and John S. Earle. 2001. “Privatization, Competition, and Reform Strategies: Theory and Evidence from Russian Enterprise Panel Data.” SITE Working Paper. Stockholm Institute of Transition Economics.
- Burkart, Mike, Denis Gromb, and Fausto Pannunzi. 1997. “Large Shareholders, Monitoring, and the Value of the Firm.” *Quarterly Journal of Economics* 112: 693–728.
- Djankov, Simeon, and Peter Murrell. 2000. *The Determinants of Enterprise Restructuring in Transition: an Assessment of the Evidence*. Washington, D.C.: World Bank.
- Earle, John S., and Saul Estrin. 1997. “After Voucher Privatization: The Structure of Corporate Ownership in Russian Manufacturing Industry.” CEPR Working Paper 1736. Centre for Economic Policy Research, London.
- . 1998. “Privatization, Competition, and Budget Constraints: Disciplining the Behavior of Russian Enterprises.” SITE Working Paper. Stockholm Institute of Transition Economics.
- Earle, John S., Roman Frydman, and Andrzej Rapaczynski. 1993. “Introduction: Privatization Policies in Eastern Europe: Diverse Routes to a Market Economy.” In John S. Earle, Roman

- Frydman, and Andrzej Rapaczynski, eds., *Privatization in the Transition to a Market Economy: Studies of Preconditions and Policies in Eastern Europe*. New York: St. Martin's Press.
- EBRD (European Bank for Reconstruction and Development). 1999. *Transition Report 1999: Ten Years of Transition*. London.
- Frydman, Roman, and Andrzej Rapaczynski. 1991. "Markets and Institutions in Large-Scale Privatization: An Approach to Economic and Social Transformation in Eastern Europe." In *Reforming Central and Eastern European Economies: Initial Results and Challenges; A World Bank Symposium*. Washington, D.C.: World Bank.
- Frydman, Roman, Cheryl Gray, Marek Hessel, and Andrzej Rapaczynski. 1999. "When Does Privatization Work? The Impact of Private Ownership on Corporate Performance in the Transition Economies." *Quarterly Journal of Economics* November: 1153–92.
- Joskow, Paul L., Richard Schmalensee, and Natalia Tsukanova. 1994. "Competition Policy in Russia During and After Privatization." *Brookings Papers on Economic Activity: Microeconomics*. 301–74.
- Kornai, Janos. 1990. *The Road to a Free Economy Shifting from a Socialist System: The Example of Hungary*. New York: W.W. Norton.
- . 1992. *The Socialist System: The Political Economy of Communism*. Princeton, N.J.: Princeton University Press.
- . 2000. "Making the Transition to Private Ownership." *Finance and Development* 37 (September): 12–13.
- Lipton, David, and Jeffrey Sachs. 1990. "Privatization in Eastern Europe: The Case of Poland." *Brookings Papers on Economic Activity: Microeconomics* 2: 293–333.
- Murrell, Peter. 1992. "Evolution in Economics and in the Economic Reform of the Centrally Planned Economies." In Christopher C. Clague and Gordon Rausser, eds., *Emerging Market Economies in Eastern Europe*. Cambridge: Blackwell.
- Shleifer, Andrei, and Daniel Treisman. 2000. *Without a Map: Political Tactics and Economic Reform in Russia*. Cambridge, Mass: MIT Press.
- Spicer, Andrew, Gerald A. McDermott, and Bruce Kogut. 2000. "Entrepreneurship and Privatization in Central Europe: The Tenuous Balance Between Destruction and Creation." *Academy of Management Review* 25 (July): 630–49.
- Stark, David, and Laszlo Bruszt. 1997. "Path Dependence and Privatization Strategies." In David Stark and Laszlo Bruszt, eds., *Postsocialist Pathways: Transforming Politics and Property in East Central Europe*. Cambridge: Cambridge University Press.
- Stiglitz, Joseph E. 1994. *Whither Socialism? The Wicksell Lectures*. Cambridge, Mass: MIT Press.
- . 2000. "Whither Reform? Ten Years of the Transition." In Boris Pleskovic and Joseph E. Stiglitz, eds., *Annual World Bank Conference on Development Economics 1999*. Washington, D.C.: World Bank.

Comment on “Ownership Structure, Legal Protections, and Corporate Governance,” by I. J. Alexander Dyck

Rafael La Porta

I. J. Alexander Dyck argues that investor protection is essential for the development of capital markets. He emphasizes that investor protection includes not only the letter of the law but also its enforcement. Enforcement requires both an efficient judiciary and the availability of good financial information that can be used to monitor insiders. La Porta and others (1997, 1998) make a similar argument and show that both legal rules and their enforcement (as measured by the efficiency of the judiciary and the quality of accounting) are important determinants of the level of investor protection and the development of capital markets. Dyck presents new evidence that suggests that firms with widely held ownership structures may simply not be viable in countries with weak investor protection. He argues that setting up ownership structures (through privatization, for example) that are inconsistent with the existing level of investor protection can be enormously wasteful. He concludes by making the case for a broad program of corporate governance reform that includes writing better laws as well as creating the institutions that will facilitate their enforcement.

Dyck’s article is a difficult one to review because I agree with so much of what is in it. Rather than disagree on minor issues, I focus on an important aspect of investor protection that Dyck does not cover at length: legal origin and its implications for legal reform. My comments are organized in three sections. I first examine the source of the differences in investor protection across countries and argue that these differences are intimately tied to the legal origin of countries. I then discuss the implications of the legal approach to corporate governance for policy reform. In the last section I present my conclusions.

Legal Origin

There are significant differences across countries in how well legal rules protect outside investors. Common law countries provide the strongest protection of outside investors, both shareholders and creditors; French civil law countries provide the

Rafael La Porta is associate professor of economics at Harvard University. These comments borrow heavily from Johnson and others (2000) and La Porta and others (2000).

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

weakest protection (La Porta and others 1998). German civil law countries are in between, although they provide stronger protection of creditors, especially secured creditors than French legal origin countries do. Countries with laws of Scandinavian origin are similar to German law countries. In general, differences between legal origins are best described by the proposition that some countries protect all outside investors better than others, not by the proposition that some countries protect shareholders and others protect creditors.

Countries differ systematically in how they enforce legal protections for outside shareholders. The quality of enforcement also has several elements, from the efficiency of the judiciary to the quality of accounting standards. The quality of enforcement is higher in richer countries, although legal rules themselves do not appear to depend on the level of economic development. But here as well legal origin matters: with the level of per capita income constant, countries with French legal origin have the lowest quality law enforcement of the four legal traditions, and countries with Scandinavian legal origin the highest.

It is unlikely that laws were written primarily in response to market pressures, since legal origins, which long preceded the development of financial markets, are highly correlated with the content of the law. Instead, legal origins appear to shape legal rules, which in turn influence financial markets. But what is special about legal origins? Why, in particular, is common law more protective of investors than civil law?

There is no consensus on the answer to these questions. It may be useful to distinguish between two broad kinds of answers: "judicial" explanations, which account for differences in legal philosophies using the organization of the legal system, and "political" explanations, which account for the differences using political history. The "judicial" explanation of why common law protects investors better runs as follows.¹ In all major legal systems courts assess the legality of transactions on the basis of two broad principles. The first is duty of care, which in this context refers to the responsibilities of corporate directors (and of controlling shareholders insofar as they also serve as directors). The duty of care, derived from the Roman concept of *mandatum*, requires directors to act the way a reasonable, prudent, rational person would act in their position. In most countries courts have implemented the duty of care using the "business judgment rule," which gives directors the benefit of the doubt unless plaintiffs demonstrate gross negligence. In the United States, for example, case law on executive compensation and takeover resistance is typically governed by the business judgment rule. Not surprisingly, these are the areas in which perceived abuse of minority shareholders is most significant in the United States.

The second general principle is the duty of loyalty, which addresses situations in which a conflict of interest arises. This duty requires that an agent not take improper advantage of the principal. In corporations this means that directors must not take advantage of shareholders or of anyone else to whom they owe the duty of loyalty (employees, "the corporation," or even "the business group"). The duty of loyalty goes beyond "rational and prudent" conduct of directors and, depending on the legal system, is governed by statutes prohibiting particular conduct or by more gen-

eral notions of fairness. In situations involving a conflict of interest, the duty of care may allow directors to take actions that favor themselves as long as they can argue that the shareholders also benefited or were not hurt too badly, at least in the long run. In contrast, the duty of loyalty may rule out such self-serving conduct.

The Belgian court case of Flambo and Barro illustrates the difference between the two duties (Wymeersch 1993). Flambo, a French firm, was the controlling shareholder in Barro, a Belgian company. Arguing that Flambo had stripped Barro of its assets, several significant minority shareholders of Barro sued Flambo, demanding judicial intervention and remedies. The plaintiffs argued that Flambo had tried to pledge Barro (the entire company) as collateral to guarantee Flambo's debt, forced Barro to acquire all the new shares issued by Flambo in a capital increase, withdrawn a substantial sum from Barro's accounts without subsequent repayment, diverted an important contract with Rank Xerox from Barro to Flambo, and used utilities belonging to Barro without paying for them.

The court barred Flambo from continuing to transfer resources from Barro without judicial review pending the election of a new board of directors for Flambo. But it did not propose any remedies for past expropriation or even a change in Barro's board. Relying on the business judgment rule, the court held that Flambo's conduct was consistent with the interest of the group as a whole. It found nothing inherently objectionable in a subsidiary supporting its parent as long as the subsidiary itself was not in danger of bankruptcy. Fairness to Barro's minority shareholders did not come up in the ruling. This and other cases illustrate that in civil law countries insiders face a very low standard of proof in cases of minority shareholder abuse; this results from the narrow application of the concept of fiduciary duty and the reliance on statutes (see Johnson and others 2000).

Civil law courts' greater reliance on statutes rather than general principles in judging conduct is not accidental. In fact, it was central to the design of the civil code by Napoleon and his jurists, who feared monarchist judges and felt that good laws should turn judges into automatons rather than interpreters of the law. This conscious choice may reflect an optimal balance between the goals of regulating the behavior of individuals through the law and controlling the behavior of those who enforce the law to ensure that its application is consistent with the goals of the social planner (Glaeser and Shleifer 2000).

Thus although the application of the law may be more predictable in civil law than in common law countries, courts in civil law countries cannot stop self-dealing transactions with a plausible business purpose, because motivated insiders can structure transactions to comply with the letter of the law. Judges in civil law countries are not allowed to apply a "smell test" to uncover unfair conduct. As a consequence, a corporate insider who finds a way to expropriate outside investors that is not explicitly forbidden by the law can proceed without fear of an adverse judicial ruling.

In contrast, legal rules in the common law system are usually made by judges based on precedents and inspired by general principles such as fiduciary duty. Judges are expected to rule on new situations by applying these general principles even when the statutes have not yet described or prohibited the specific conduct under

judicial review. Judges apply special scrutiny to transactions that "smell bad." The expansion of legal precedents to additional violations of fiduciary duty and the fear of such expansion limit expropriation by insiders in common law countries. From this perspective, the vague fiduciary duty principles of common law are more protective of investors than the "bright line" rules of civil law, which can often be circumvented by sufficiently imaginative insiders.

The judicial perspective on the differences between legal origins is incomplete. It requires a further assumption that judges have an inclination to protect outside investors rather than insiders. But it is easy to imagine that judges in common law countries would use their discretion to narrow the interpretation of fiduciary duty and to sanction rather than prohibit expropriation. In principle, common law judges could also use their discretion to serve political interests, especially when outside investors obstruct the government's goals. To explain investor protection, it is not enough to focus on judicial power; a political and historical analysis of judicial objectives is required. Important political and historical differences between mother countries shape their laws. This is not to say that laws never change but to suggest that history has persistent effects. How so?

La Porta and others (2000) argue that an important historical factor shaping laws is the state's greater role in regulating business in civil law countries than in common law countries. One element of this view, suggested by Finer (1997) and other historians, is the difference across European states in the relative power of the king and property owners. In England from the 17th century on (and arguably before), the crown partially lost control of the courts, which came under the influence of Parliament and the property owners who dominated it. As a consequence, common law evolved to protect private property against the crown. Over time courts extended the protection of property to investors.

In France and Germany, in contrast, the parliaments never dominated the kings, and the state dominated the courts and property owners. Commercial codes were adopted only in the 19th century, by the two great state builders, Napoleon and Bismarck, to enable the state to better regulate economic activity. As the law evolved, the dominance of the state translated into the more political conception of the corporation and the more limited rights of investors in dealing with the politically connected families that control firms. After all, the state was not about to surrender its power over business to financiers.² Relative to courts in England, courts in civil law countries were more dependent on the government and less likely to take the side of investors in disputes with the government or with firms close to it. As a consequence of these divergent political histories, civil law developed to become less protective of investors than did common law.

Recent research supports the proposition that civil law is associated with greater government interference in economic activity and weaker protection of private property. La Porta and others (1999) examine the determinants of government performance in a large number of countries. To measure government interventionism, they consider proxies for the quality of regulation, the prevalence of corruption and red tape, and bureaucratic delays. They find that as a general rule, governments of

civil law countries, particularly French civil law countries, are more interventionist than those of common law countries. The inferior protection of the rights of outside investors in civil law countries may be one manifestation of this general phenomenon. This evidence thus provides some support for interpreting the differences between legal origins in light of political history.

Possibilities for Reform

In the past decade the reform of corporate governance has attracted interest in Asia, Latin America, and Western and Eastern Europe. Not much appears to have been done, however, although interest remains high.

To discuss any reform, it is important to start with its goals. In most countries the objective of corporate governance reform is to protect the rights of outside investors, both shareholders and creditors. As the evidence surveyed by Dyck shows, such reform would expand financial markets, facilitate external financing of new firms, reduce the concentration of ownership, and improve the efficiency of investment allocation.

What can be done to achieve this goal, and what are the obstacles to doing so? To organize this discussion, it is useful to follow Coffee (1999b) in distinguishing between legal and functional convergence. Legal convergence refers to the changes in rules and enforcement mechanisms toward some desirable standard. To achieve legal convergence to effective investor protection, most countries need extensive legal, regulatory, and judicial reform. Functional convergence refers to more decentralized, market-based changes; these do not require legal reform but still bring more firms and assets under the umbrella of effective legal protection of investors.

For most countries, improving investor protection requires rather radical changes in the legal system. Security, company, and bankruptcy laws generally need to be amended. As Dyck points out, there is no reason to think that the list of legal protections of investors studied by La Porta and others (1998) is either necessary or sufficient for such reforms. Moreover, the regulatory and judicial mechanisms of enforcing shareholder and creditor rights need to be radically improved. The evidence on the importance of the historically determined legal origin in shaping investor rights—which could be thought of as a proxy for the law's general stance toward outside investors—suggests that many rules and institutions need to be changed simultaneously in a country with poor investor protection to bring them up to best practice.

The politics of such change have proved brutal. Effective legal reform runs into tremendous political obstacles. Interventionist governments may not be interested in ceding to financiers the control they currently have over large corporations. In addition, families that control large corporations strongly resist improvements in investor protection. From their point of view, an improvement in the rights of outside investors is first and foremost a reduction in the value of control, because it reduces opportunities for expropriation. Legal reform may increase the total value of these firms as expropriation declines and investors finance new projects on more attractive terms.

Still, the first-order effect on insiders is a massive redistribution of wealth from them to outside investors. No wonder then that controlling shareholders—from Latin America to Western and Eastern Europe to Asia—oppose legal reform. As Hellwig (1999) describes the attitude of German industrialists to corporate governance reform, "You can wipe my face but do not get it wet." The German government has opposed transparency and other governance reforms in the European Union.

There is another reason why insiders in major firms oppose corporate governance reform and the expansion of capital markets. Under the status quo, existing firms can finance their investment projects internally (Mayer 1988) or through captive or closely connected banks. In fact, La Porta and others (1997) show that the lion's share of credit in countries with poor investor protection goes to the largest firms. As a consequence, large firms obtain not only the finance they need but also the political influence that comes with access to such finance as well as protection from the competition that could develop if smaller firms could also raise external capital. When new entrepreneurs have good projects, they often have to come to existing firms for capital. For the insiders, poor corporate governance delivers not only secure finance but also secure politics and markets. They thus have an interest in keeping the system as it is.

Consistent with the dominance of interest group politics, successful reforms have occurred only when special interests have been destroyed or appeased. In this respect, corporate governance reform is no different from most other kinds of reform in developing or industrial countries (Hirschman 1963). One example of significant legal reform of corporate governance occurred in Japan after World War II, where General Douglas MacArthur, aided by attorneys from Chicago and an occupying army, introduced a company law based on Illinois law (Ramseyer and Nakazato 1999). Another example is the introduction of regulations for securities markets in the United States in 1933–34, immediately after President Franklin D. Roosevelt's election, in the middle of the Great Depression. Although these regulations have been criticized as ineffective (Stigler 1964), there is little doubt that they have substantially increased corporate disclosure in the United States. Still another example, though one in which success has been more limited, is bankruptcy reform in East Asia following the 1997 crisis. These examples suggest that opportunities for corporate governance reform do arise, but under special circumstances, which should not be wasted.

Unfortunately, opportunities have often been wasted, because our understanding of the principles of reform of investor protection remains limited. There is no checklist of what needs to be done. However, recent research suggests some tentative principles.

Perhaps the first such principle is that legal rules do matter; it is not just the stance of the law or the political sentiment that shapes financial markets. One illustration of this principle is the Neuer Markt in Germany, a subexchange that the Frankfurt Stock Exchange created especially for new firms wishing to go public (Johnson 1999). Because the Neuer Markt operates in Germany, the corporate law, securities law, and other basic laws and regulations applied to the companies listed there are German rules. The politics are German as well. The Deutsche Bourse has ordered that companies wishing to list on the Neuer Markt must comply with international accounting standards, including more stringent disclosure requirements than those

applicable to firms already listed on the main exchange. The new listing venue—with its greater restrictions on entrepreneurs—has sharply accelerated the pace of initial public offerings in Germany. The captains of German industry have accepted it because it has not affected their firms. This example thus points to one possible strategy for overcoming political opposition to reform.

A second principle is that the enforcement of legal rules is deeply connected with the rules themselves. The strategy for reform is not to create an ideal set of rules and then see how they can be enforced but to enact rules that can be enforced with the existing enforcement structure. A famous example of success with such a policy is the securities legislation enacted in the United States in 1933–34. The creator of these laws, James Landis, a professor of administrative law, was deeply focused on enforcement (McGraw 1984). His strategy for forcing corporate disclosure was to demand independent audits of firms' accounts and to make the accountants liable for their reports. Although accounting firms initially resisted these reforms, they soon appreciated the increase in demand for their services. The accounting profession thus became an independent, private force in ensuring compliance with disclosure regulations with fairly limited government involvement. At the time at least, it would have been impossible for the government to enforce disclosure regulations on its own without creating incentives for private self-regulation. The principle of bringing private intermediaries into the enforcement of securities regulations has since been followed by several countries, including Germany and Poland, in regulating their financial markets.

The importance of enforcement is also illustrated by the still unsuccessful reform of bankruptcy procedures in East Asia. In general, improving bankruptcy procedures is harder than improving shareholder rights, because different types of creditors, unlike different noncontrolling shareholders, have different objectives. Senior creditors, especially secured senior creditors, prefer rapid liquidation of bankrupt firms. Junior creditors and shareholders, whose claims are less secure, may prefer more orderly liquidation or even reorganization. As a result of this conflict, most countries have opted for rather slow, reorganization-focused bankruptcy schemes rather than liquidations (Hart 1999).

The trouble is, bankruptcy procedures almost inevitably rely on significant adjudication by the courts. Yet courts in many countries are reluctant to be too active in matters as political and complicated as closing or liquidating companies. In the aftermath of the Asian crisis several East Asian countries, including Indonesia, the Republic of Korea, and Thailand, reformed their bankruptcy laws. Few companies have been taken through the bankruptcy process, however, largely because the courts, which are politicized and unready to adopt the new procedures, throw out most creditor applications on technicalities.

A third—and potentially more controversial—principle of successful reform is that government regulation of financial markets may be useful when court enforcement of private contracts or even of government laws cannot be relied on. An example of how regulation can work when courts are far from perfect comes from securities law reform in Poland and the Czech Republic, two transition economies whose judiciaries were viewed as ineffective. In the early 1990s the Polish government introduced

a tough securities law focused on protecting shareholders. Like U.S. securities law, the Polish regulations focused on significant disclosure by issuers and listed firms and the legal accountability of intermediaries for the accuracy of this information. The law also provided for the creation of a powerful securities and exchange commission with significant enforcement powers that did not require reliance on the courts. This reform was followed by remarkable development of the Polish stock market, with both new and listed companies raising equity in the market.

In contrast, the Czech government chose neither to introduce tough securities laws nor to create a powerful market regulator at the time of privatization. Perhaps as a consequence, the Czech markets have been plagued by massive expropriation of minority shareholders—the “tunneling” of assets from both firms and mutual funds. In contrast to the Polish market, the Czech market stagnated, with hundreds of companies delisted and virtually no public equity financing by firms (see Coffee 1999b; Pistor 1999; and Glaeser, Johnson, and Shleifer forthcoming). The comparison of Poland and the Czech Republic is an almost perfect experiment, because the two countries share roughly similar incomes, economic policies, and judiciary quality. In these circumstances regulation of the stock market and listed firms in Poland, with its focus on investor protection, appeared to play a beneficial role.

The successful regulation of the U.S. securities markets, the Polish financial markets, and the Neuer Markt in Germany shares a common element—the regulatory insistence on extensive disclosure of financial information by the issuers, the accuracy of which is enforced by the legally liable financial intermediaries. Although such disclosure is not sufficient by itself—without the right of shareholders to act on it—it appears to be a key element of shareholder protection.

Although legal reform has been slow and halting in most countries, functional convergence may play a role in improving investor protection. The liberalization of capital markets in many countries has increased not only the flow of foreign investment into them, as Henry (2000) and Stulz (1999) document, but also the economic and political pressure to create financial instruments acceptable to foreign investors. These pressures give rise to several forms of functional convergence. Most obviously, if contracts are enforced well, companies in unprotective legal regimes can offer their investors customized contracts, such as corporate charters, with greater rights than the law generally provides. This strategy relies on perhaps a greater faith in the contracting capacity of investors and courts than is warranted and ignores the empirically clear public good benefit of standard rules.

A more promising approach is for companies to opt into the more investor-friendly legal regimes. One way of doing so is to list a company's securities on an exchange that protects minority shareholders through disclosure or other means. This is done by the many companies that list their shares as American depositary receipts (ADRs) in New York. But such listing imposes only limited constraints on insiders: although it improves disclosure, it typically does not give minority shareholders any effective rights.

A related mechanism for opting into a more protective legal regime is acquisition by a company already operating in such a regime. When a British company acquires a Greek company, the possibilities for legal expropriation of investors diminish. In

a friendly deal such as this, the controlling shareholders of the Greek company can be compensated for the lost private benefits of control, making them more likely to go along. By replacing the wasteful expropriation with publicly shared profits and dividends, such acquisitions enhance efficiency.

It is important to recognize the limitations of functional convergence, particularly in creditor rights. Assets generally remain under the jurisdiction of the laws of the country in which they are located. Without bankruptcy reform, opt-in mechanisms are unlikely to address the legal problems faced by domestic, and especially foreign, creditors. Despite the benefits of opting into the more protective legal regime for external finance, then, this mechanism is unlikely to fully replace bona fide legal reform. Slow and difficult as it is, real legal reform needs to take place in many countries.

Conclusion

Leaving financial markets alone is not a good way to encourage them. Financial markets need some regulation protecting outside investors—whether by courts, government agencies, or market participants themselves. The evidence suggests that civil law countries are less protective of minority shareholders than are common law countries, in terms of both laws and their enforcement. Moreover, important differences in the ways in which countries from different legal traditions apply laws may facilitate minority shareholder expropriation in civil law countries. The link between investor protection and legal origin is probably not accidental. Strong investor protection may be a particularly important manifestation of the greater security of property rights against political interference in some countries.

Improving the regulation of financial markets is a difficult task for two reasons. First, the nature of investor protection, and of the regulation of financial markets more generally, is deeply rooted in the legal structure of each country and in the origin of its laws. Reform on the margin may not achieve the reformer's goals. Second, the existing corporate governance arrangements benefit both the politicians and the entrenched economic interests, such as the families that manage the largest firms in most countries. Corporate governance reform must circumvent opposition by these interests.

Despite these difficulties, investor protection reform can bring significant benefits, and it is politically feasible in some circumstances. It can take the form of opting into more protective legal regimes or introducing more radical change in the legal structure. The integration of world capital markets makes such reforms more likely today than they have been in decades.

Notes

1. This explanation is consistent with David and Brierley (1985) and Coffee (1999a)
2. According to Cameron (1961), France had a lively stock market in the 19th century. Nearly all firms listed on it, however, benefited from government concessions, investment, ownership, subsidies, and protection, as well as outright guarantees to investors in many cases.

References

- Cameron, Rondo E. 1961. *France and the Economic Development of Europe*. Princeton, N.J.: Princeton University Press.
- Coffee, John C., Jr. 1999a. "Privatization and Corporate Governance: The Lessons from Securities Market Failure." *Journal of Corporation Law* 25: 1–39.
- . 1999b. "The Future as History: The Prospects for Global Convergence in Corporate Governance and Its Implications." *Northwestern Law Review* 93: 631–707.
- David, René, and John Brierley. 1985. *Major Legal Systems in the World Today*. London: Stevens and Sons.
- Finer, Samuel. 1997. *The History of Government*. Cambridge: Cambridge University Press.
- Glaeser, Edward, and Andrei Shleifer. 2000. "Legal Origins." Harvard University, Department of Economics, Cambridge, Mass.
- Glaeser, Edward, Simon Johnson, and Andrei Shleifer. Forthcoming. "Coase versus the Coasians." *Quarterly Journal of Economics*.
- Hart, Oliver. 1999. "Different Approaches to Bankruptcy." Harvard University, Department of Economics, Cambridge, Mass.
- Hellwig, Martin. 1999. "On the Economics and Politics of Corporate Finance and Corporate Control." University of Mannheim, Department of Economics, Germany.
- Henry, Peter Blair. 2000. "Do Stock Market Liberalizations Cause Investment Booms?" *Journal of Financial Economics* 58 (1): 301–34.
- Hirschman, Albert. 1963. *Journeys toward Progress*. New York: Twentieth Century Fund.
- Johnson, Simon. 1999. "Do Legal Rules Matter? Evidence from Germany's Neuer Markt." Massachusetts Institute of Technology, Sloan School of Management, Cambridge, Mass.
- Johnson, Simon, Rafael La Porta, Florencio López-de-Silanes, and Andrei Shleifer. 2000. "Tunnelling." *American Economic Review: Articles and Proceedings* 90 (2): 22–27.
- La Porta, Rafael, Florencio López-de-Silanes, Andrei Shleifer, and Robert Vishny. 1997. "Legal Determinants of External Finance." *Journal of Finance* 52: 1131–50.
- . 1998. "Law and Finance." *Journal of Political Economy* 106: 1113–55.
- . 1999. "The Quality of Government." *Journal of Law, Economics and Organization* 15: 222–79.
- . 2000. "Investor Protection and Corporate Governance." *Journal of Financial Economics* 58 (1): 3–28.
- Mayer, Colin. 1988. "New Issues in Corporate Finance." *European Economic Review* 32: 1167–88.
- McGraw, Thomas K. 1984. *Prophets of Regulation*. Cambridge, Mass.: Harvard University Press.
- Pistor, Katharina. 1999. "Law as a Determinant of Equity Market Development." Harvard University, Kennedy School of Government, Cambridge, Mass.
- Ramseyer, Mark, and Minoru Nakazato. 1999. *Japanese Law: An Economic Approach*. Chicago: University of Chicago Press.
- Stigler, George. 1964. "Public Regulation of the Securities Market." *Journal of Business* 37: 117–42.
- Stulz, René M. 1999. "International Portfolio Flows and Securities Markets." Working Paper 99-3. Ohio State University, Fisher College of Business, Columbus.
- Wymeersch, Eddie. 1993. "Groups of Companies in the EEC: A Survey Report to the Commission on the Law Relating to Corporate Groups in Various Member States." Berlin: W. De Gruyter.

*Social Security,
Public and
Private Savings*

The Benefits of Flexible Funding: Implications for Pension Reform in an Uncertain World

J. Michael Orszag and Peter R. Orszag

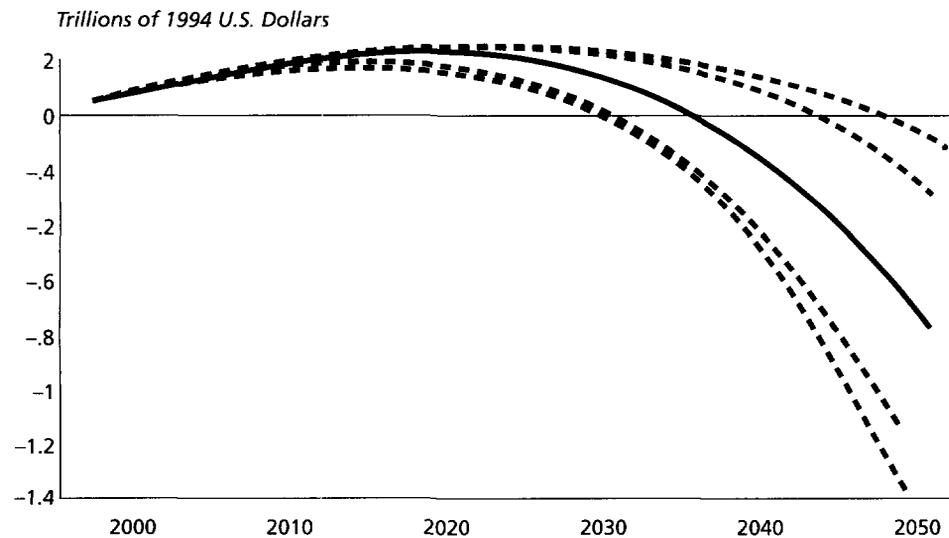
Aging populations are expected to put increasing strains on government budgets across the globe, but the extent of the pressure is unclear because of the substantial uncertainty surrounding estimates of pension costs several decades into the future. This article examines the implications of uncertainty over future pension costs for current decisions about how to undertake pension reforms, especially decisions involving prefunding. Prefunding can be thought of as an investment that initially entails some costs (during the transition to the new prefunded system) and then offers long-run benefits. In the face of uncertainty over long-run benefits, systems with more flexibility in the level of funding may be preferable to systems with less flexibility, because they allow policy to adapt more quickly should circumstances change.

As the proportion of the world's population 60 and older increases—from 9 percent in 1990 to 16 percent in 2030, according to World Bank (1994) projections—aging populations are expected to put increasing strains on government budgets across the globe (Lee and Skinner 1999). Substantial uncertainty surrounds estimates of pension costs several decades into the future, however. Actuaries for the U.S. Social Security Administration traditionally prepare three estimates: low cost, intermediate cost, and high cost. The intermediate estimate, commonly cited in the press, suggests a nontrivial actuarial deficit, but the low-cost estimate shows a small actuarial surplus over the next 75 years at the current payroll tax rate (U.S. Social Security Administration 2000). Many observers have objected to the manner in which Social Security actuaries reflect parameter uncertainty (Lee and Skinner 1999). But more sophisticated methodologies also show substantial uncertainty (figure 1).

J. Michael Orszag is lecturer in economics at Birkbeck College, in London. Peter R. Orszag is president of Sebago Associates, in Belmont, California. The authors are grateful to Stephen Diacon, Louise Fox, Robert Holzmann, Joseph E. Stiglitz, Diane Whitmore, participants at the Annual World Bank Conference on Development Economics 2000, and two anonymous reviewers of initial drafts for helpful discussions and comments.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

Figure 1. Central Estimate and Confidence Bands for U.S. Social Security Trust Fund Balance, 2000–50

Note: The confidence bands present the 67 percent and 95 percent confidence intervals. That is, the trust fund is projected to remain within the outermost set of lines with 95 percent confidence.
Source: Lee and Tuljapurkar 1999.

As Lee and Tuljapurkar (1998) conclude, “Rational planning for the next century must somehow take into account not just our best guesses about the future but also our best assessments of the uncertainty surrounding these guesses” (p. 393).

We examine the implications of uncertainty about future pension costs for current decisions on pension reforms. In particular, we show the importance of undertaking prefunding in a way that allows flexibility, so that the level of prefunding can be adjusted to reflect new developments.

The model we use adopts a “real options” perspective on pension reform. In this approach pension reform can be thought of as an investment that initially entails some costs (during the transition to the new system) and then offers long-run benefits. In the face of uncertainty over the long-run benefits, a key question is whether the costs can be recaptured. Systems that allow funding to be scaled up or down more easily may be preferable to less flexible alternatives, because flexibility allows policymakers to alter policy more quickly should circumstances change and thus to partially recapture the costs that had been incurred by funding in the first place. Under an inflexible system the costs of funding are much more difficult to recapture, even if it becomes obvious that the funding had been unnecessary.

An Illustrative Example

Consider a country with a pay-as-you-go pension system for which actuarial and economic analysis indicates that the social welfare cost will rise to 2 percent of

GDP next period. As just one motivating example for the social cost, assume that payroll taxes distort labor market choices and that the pay-as-you-go pension system is financed through payroll taxes that are higher than they would be in its absence.¹

The social welfare costs are highly uncertain, however, because of uncertainties over fertility, mortality, productivity, and other economic variables, uncertainties that affect the cost rate of the pay-as-you-go pension system and therefore the social welfare costs from the distortions caused by payroll taxes. The social costs could be 1 percent of GDP or 3 percent of GDP, with 50 percent probability of each; the expected value is thus 2 percent of GDP. We assume that the social costs next period will be perpetuated into the future indefinitely.

By raising national saving today, the government can increase GDP next period and reduce the social welfare costs of the pension system as a percentage of GDP. Let us assume, again for simplicity, that broad prefunding would reduce the cost of the pension system to 1 percent of GDP next period and that there would then be no uncertainty over the welfare cost from period to period. (We use “broad prefunding” in the sense of Orszag and Stiglitz (2001) to mean an increase in national saving. Pension reforms that do not engender broad prefunding are unlikely to have any significant macroeconomic benefits.)

The benefit of broad prefunding is that it reduces the expected tax rate and therefore (in this simple model) the labor market distortion in the long run. Indeed, Feldstein and Samwick (1998) highlight this reduction in tax rates as a key advantage of funding.²

Such broad prefunding has a cost, however: the additional revenue necessary to produce it. In the absence of nondistortionary taxes, raising the additional revenue necessary to prefund the pension system will impose deadweight losses on the economy during the initial phase. (Broad prefunding requires some type of additional revenue; pure debt financing is not consistent with an increase in national saving. The source of additional revenue could be chosen to minimize the additional deadweight losses, but absent the availability of nondistortionary taxes at the margin, some additional deadweight loss is necessary.)

We assume that the temporary cost of increased distortions in the short run is equal to a one-time cost of 10 percent of GDP, which includes both the cost of paying down the implicit debt under the pay-as-you-go system and the cost of financing the pension system in the period during which the transition cost is borne. The unfunded implicit debt of the pay-as-you-go system is a stock that could be many multiples of GDP. Bravo and Uthoff (1999), for example, calculate implicit pension debts of more than 200 percent of GDP for Brazil in 1990 and more than 300 percent of GDP for Argentina in 1990. The social welfare costs of paying down the debt may thus be substantial in the short run, depending on the form of taxation used to do so.

Now consider a social planner trying to decide whether to prefund the pension system. At a real social discount rate of 5 percent, the expected present value (expressed relative to GDP) of the prefunding is:

$$(1) \quad -10 + \sum_{t=1} \frac{2-1}{1.05^t} = -10 + 20 = 10$$

The prefunding thus appears to be socially beneficial. Several points about this example are noteworthy:

- The existing pay-as-you-go pension system is expected to become more costly over time, but there is uncertainty about how costly it will be.
- Prefunding is socially beneficial, but focusing solely on the expected differential between the pay-as-you-go system and the prefunded system in the long run is misleading. A full analysis must take into account the transition costs of moving from the pay-as-you-go to the prefunded system.
- The analysis is undertaken in terms of social welfare costs, not financial costs. The net present value of a pay-as-you-go system across all generations is always zero (Geanakoplos, Mitchell, and Zeldes 1999). This zero net present value condition implies that there is no free lunch from prefunding in net present value terms, but it does not require that the impact on social welfare be zero. Indeed, most economists support prefunding because they believe that it boosts social welfare. The social welfare implications depend on a broader array of considerations—including intergenerational and intragenerational redistribution concerns—than the financial zero net present value condition might suggest (Orszag and Stiglitz 2001). We assume in this example that the prefunding is socially beneficial, consistent with the beliefs of many economists.

This example seems to capture the essence of the argument for prefunding—that the existing pay-as-you-go system is expected to become more costly, that prefunding can reduce the long-run cost, and that prefunding improves social welfare. The example conforms generally to the argument often put forward in favor of prefunding. Feldstein and Samwick (1998), for example, note that the long-run reduction in tax rates and deadweight losses from prefunding follows only after an extended period (23 years in their base-case scenario) in which overall contribution rates and deadweight losses are higher than under the existing system.

If prefunding is irreversible, however, an alternative strategy may be preferable: waiting to see how the demographic and economic uncertainty resolves itself and then prefunding only if the “bad” case arises. In particular, consider the strategy of waiting until next period and prefunding only if the pay-as-you-go welfare cost turns out to be 3 percent of GDP. (If the welfare cost instead turns out to be 1 percent of GDP, the pay-as-you-go system would be perpetuated.) The expected net present value of that strategy is:

$$(2) \quad (0.5) - \frac{10}{1.05} + \sum_{t=2} \frac{3-1}{1.05^t} = 0.5(-9.5 + 38.1) = 14.3$$

This strategy thus produces a higher expected value than the straightforward funding strategy. This result occurs because the strategy avoids paying the transition costs in the scenario in which the social costs of the pay-as-you-go system turn out to be unexpectedly low. In that case, having paid the transition costs would (in this example) have been a mistake *ex post*.

The higher expected value produced in equation 2 relative to equation 1 could be seen as supporting the arguments by opponents of prefunding, who suggest that pension projections are highly uncertain and that it is therefore not worth incurring current costs to eliminate a projected imbalance. Dean Baker, of the Center for Economic and Policy Research, has stated that “it’s hard to see what the benefits are of acting now. We want to plan ahead but want to have some idea of what we’re planning for.... The idea that we have to do something now because 40 years from now the U.S. Social Security Trust Fund will be out of balance doesn’t make sense” (cited in Pianin 2000, p. 4). He argues that internationally the expected payroll tax increase under existing pay-as-you-go systems is not as burdensome as often depicted (Baker 1999).

An alternative strategy, however, dominates the Baker approach in this model. In particular, consider a strategy that could be called “conditional prefunding.”³ Under this strategy prefunding is undertaken today but reversed tomorrow if the welfare costs under the pay-as-you-go system turn out to be 1 percent of GDP rather than 3 percent. Assuming that the 10 percent of GDP can be recaptured next period if the “good” case materializes, the net present value of that strategy is then:

$$(3) \quad (0.5) - 10 + \frac{10}{1.05} + (0.5) - 10 + \sum_{t=1} \frac{3-1}{1.05^t} = 0.5(-.48) + 0.5(-10 + 40) = 14.8$$

This conditional prefunding approach has the highest net present value of all. The reason is that it does not wait to prefund, as the Baker strategy does, when prefunding turns out to be socially beneficial and, unlike the irreversible prefunding approach, it eliminates the cost of prefunding when the prefunding turns out to have been unnecessary.

The implication is that prefunding can be socially beneficial but that systems in which prefunding can be adjusted as information changes may be superior to systems in which the level of prefunding is irreversible. As an extreme example, assume that prefunding occurs today but that tomorrow the economy becomes dynamically inefficient. Given the changed situation, the ability to reverse the prefunding could be valuable.

The examples above are predicated on full funding of the pension system. In reality, partial prefunding is more likely than either no funding or full funding. In a fully funded system the risks are asymmetric: Assuming that overfunding is precluded, the only possible adjustment is to reduce the funding level. Therefore, assuming that funding has been undertaken, the only relevant risks are ones that make undoing

some of the funding more attractive. In a partially prefunded system, however, the risks are more symmetric: shocks may occur that warrant an increase or decrease in the funding level. Nonetheless, the key point carries across: flexibility to adjust the funding level is essential for responding appropriately to a pension challenge involving an expected but highly uncertain gain from a given level of funding.

The upshot of both the partially prefunded and the fully funded scenarios is that it is important to investigate precisely what makes one approach more or less flexible than another. In the context of pension reform “reversibility” means the ability to undo some of the funding—either by reducing the contribution rate below the funded path or by raising benefits above that path, thus reintroducing a pay-as-you-go component to the system.

The simple model described here clearly makes a number of restrictive assumptions. One of the most important may be that the funding level is adjusted solely in line with what a rational social planner would do. To the extent that flexible approaches to prefunding also allow such funding to be reversed for non-economic reasons—a phenomenon not addressed in the model—the argument in favor of flexible approaches may be weakened. But even in that case a tradeoff will exist between “socially rational” and “political economy” reversals.

A Deterministic Model

We now present a stylized model that captures the essential elements in the debate over a move toward fully funded private accounts. This section presents the deterministic model; the next section extends the model to include some uncertainty over the parameters. The intuitive, deterministic model provided here may be helpful in understanding the results presented later in the article.

Pay-As-You-Go System

Consider a simple model of a pure unfunded (pay-as-you-go) pension system in which:

- f is the growth rate of the working population of a given age (related to fertility and migration).
- d is the death rate (deaths as a percentage of the population).
- N is the length of the working life.
- t is the rate of the payroll tax used to finance benefits under the pension system.
- p is the replacement rate (pension income as a percentage of previous wages).
- g is the growth rate of real wages.

Pensions are held constant in real terms once in payment. The number of workers working at time t is:

$$(4) \quad \tilde{N}(t) = \int_0^N e^{f(t-s)} ds = e^{ft} \frac{1 - e^{-fN}}{f}$$

We assume no deaths during the working life, implying that the number of new retirees in any given year is e^{ft} . The number of retirees alive at any point in time is then:

$$(5) \quad R(t) = \int_0^{\infty} \exp(-ds) \exp(f(t-s)) ds = e^{ft} \frac{1}{d-f}$$

The elderly dependency ratio (the number of elderly people divided by the size of the working population) is thus:

$$(6) \quad \frac{R(t)}{\tilde{N}(t)} = \frac{e^{ft} \frac{1}{d-f}}{e^{ft} \frac{1-e^{-fN}}{f}} = \frac{f}{(d-f)(1-e^{-fN})}$$

For example, if $d = 0.06$, $N = 50$, and $f = 0.01$, the dependency ratio is about 0.5.

Since pensions in payment are assumed to be indexed to prices, not wages, the total pension bill as a fraction of wages at time t will be:

$$(7) \quad pe^{-fN} \int_0^{\infty} e^{-(d+g)s} e^{f(t-s)} ds$$

or

$$(8) \quad e^{ft} e^{-fN} \frac{p}{g+f+d}$$

In this case the social security tax rate is:

$$(9) \quad \tau = \frac{p}{g+d+d} \frac{f}{1-e^{-fN}} e^{-fN}$$

Funded System

We assume that our hypothetical country has a pay-as-you-go system as described above and is considering a reform in which a fully funded system would ultimately replace the pay-as-you-go system. The funding could occur either through a public trust fund or through privately managed accounts. For now we assume that the two approaches are equivalent. In the next section, when we introduce uncertainty, we explore potential differences between them.

We assume that the replacement rate under the funded system is designed to be equal to the replacement rate under the pay-as-you-go system (p). Under the funded system the contribution rate, c , would thus be defined implicitly by:

$$(10) \quad c \int_0^N e^{g(s-N)} e^{(r-k)(N-s)} ds = p \int_N^{\infty} e^{-(r-k)(s-N)} e^{-d(s-N)} ds$$

where r is the gross rate of return earned on the account (and annuity, after retirement) and k is the reduction in yield from administrative charges on the account (and annuity, after retirement). The contribution rate can then be expressed as:

$$(11) \quad c = \frac{p}{d + (r - k)} e^{-(r-k-g)N} \frac{r - k - g}{1 - e^{-(g-(r-k))N}}$$

Note that $c < \tau$ if:

$$\frac{1}{d + (r - k)} e^{-(r-k-g)N} \frac{r - k - g}{1 - e^{-(g-(r-k))N}} < \frac{1}{g + f + d} \frac{f}{1 - e^{-fN}}.$$

The system thus illustrates the Aaron (1966) rule, which states that assuming that $k = 0$, the contribution rate under the funded system is equal to the contribution rate under the pay-as-you-go system if $f + g = r$.

Transition to a Funded System

To examine the labor market distortions caused by the change in the contribution rate in the transition to the funded system, it is necessary to specify some transition path. For simplicity, we adopt an exponential path, in which the share of retirement income provided by the funded system (rather than the pay-as-you-go system) rises exponentially. The total replacement rate (from the funded and pay-as-you-go systems together) remains constant throughout the reform process. The exponential assumption simplifies the model, but it also implies that the pay-as-you-go system never completely disappears (instead, it becomes arbitrarily small).

During the transition the total replacement rate, p , for each cohort comprises two components—the part financed by the pay-as-you-go system and the part financed by the funded system:

$$(12) \quad p(Y, s) = p_f(Y, s) + p_u(Y, s)$$

where Y is the number of years since the reform, s is the number of years since the cohort's retirement, p_f is the replacement rate from the funded component of the system, and p_u is the replacement rate from the pay-as-you-go component.

The total payments to retirees (in terms of the average wage) from the pay-as-you-go component, $D_t(Y)$, where t is the year, are then given by:

$$(13) \quad D_t(Y) = \int_0^{\infty} p_u(t,s) e^{f(t-N-s)} e^{-ds} e^{-gs} ds.$$

If $Y = 0$, so that the pay-as-you-go component accounts for the entire pension, then:

$$(14) \quad p_u = \bar{p} \rightarrow D_t(0) = \frac{\bar{p} e^{f(t-N)}}{f + d + g}.$$

The exponential transition that we adopt in this model involves pay-as-you-go benefits of:

$$(15) \quad P_u(Y,s) = \begin{cases} \bar{p} & \text{if } s \geq Y \\ \bar{p} e^{\mu(Y-s)} & \text{if } s < Y \end{cases}.$$

In other words, cohorts that retire before the reform (implying that s is greater than or equal to Y) receive their entire pension from the pay-as-you-go component. Later cohorts receive a smaller and smaller share from the pay-as-you-go component. The speed with which the pay-as-you-go component declines is governed by μ .

Plugging equation 15 into equation 13, we see that the total cost of the pay-as-you-go component in year t is:

$$(16) \quad D_t(Y) = \int_0^Y \bar{p} e^{-\mu(Y-s)} e^{f(t-N)} e^{-(f+d+g)s} ds + \int_Y^{\infty} \bar{p} e^{f(t-N)} e^{-(f+d+g)s} ds$$

or:

$$(17) \quad D_t(Y) = \bar{p} e^{f(t-N)} \left[e^{-\mu Y} \frac{1 - e^{-(f+d+g-\mu)Y}}{f + d + g - \mu} + \frac{e^{-(f+d+g)Y}}{f + d + g} \right]$$

The pay-as-you-go contribution rate Y years after the reform is therefore:

$$(18) \quad \tau(Y) = \frac{D_t(Y)}{\dot{N}(t)} = \frac{\bar{f}\bar{p}e^{-fN}}{1 - e^{-fN}} \left(e^{-\mu Y} \frac{1 - e^{-(f+d+g)Y}}{f + d + g - \mu} + \frac{e^{-(f+d+g)Y}}{f + d + g} \right).$$

Together, the pay-as-you-go and funded components must produce an overall replacement rate of \bar{p} . Since the pay-as-you-go component provides a benefit of

$$\bar{p}e^{-\mu(Y+N-x)}$$

for the cohort age x years at time t , the cohort must contribute sufficiently to accumulate a fund providing a replacement rate equal to

$$\bar{p}[1 - e^{-\mu(Y+N-x)}].$$

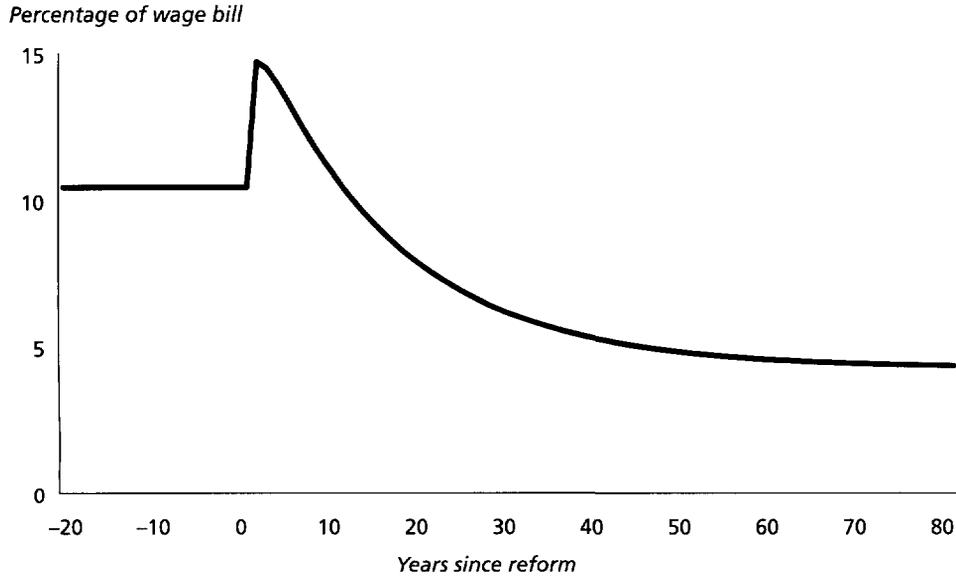
The contribution rate required for the funded component is then:

$$(19) \quad c(z, Y) = \frac{\bar{p}[1 - e^{-\mu(Y+N-z)}]e^{-(r-k-g)N}}{d + r - k} \left\{ \frac{r - k - g}{1 - e^{(g-(r-k))N}} \right\}$$

where z is the cohort's age at the time of the reform. The total contribution rate for the funded and pay-as-you-go components of the system Y years after the reform for this cohort is therefore:

$$(20) \quad c(z, Y) + \tau(Y) = \frac{\bar{p}[1 - e^{-\mu(Y+N-z)}]e^{-(r-k-g)N}}{d + r - k} \left\{ \frac{r - k - g}{1 - e^{(g-(r-k))N}} \right\} + \frac{\bar{f}\bar{p}e^{-fN}}{1 - e^{-fN}} \left[e^{-\mu Y} \frac{1 - e^{-(f+d+g-\mu)Y}}{f + d + g - \mu} + \frac{e^{-(f+d+g)Y}}{f + d + g} \right].$$

To obtain the aggregate contribution rate, $T(Y)$, which reflects the combined contribution rate as a percentage of total payroll, we weight equation 20 by the size of each cohort and integrate. Assuming that $f = 0.015$, $g = 0.015$, $d = 0.04$, $N = 40$, $p = 0.4$, $r = 0.06$, $k = 0.005$, and $\mu = 0.75$, the total contribution rate declines as a result of the reform (figure 2). Two observations are worth noting. First, the steady-state pay-as-you-go contribution rate (before the reform) is significantly higher than the steady-state funded contribution rate (apparent as Y increases and the pay-as-you-go component approaches zero). Second, to obtain the lower contribution rate in the long run, a transition is necessary in which the total contribution rate first exceeds the pay-as-you-go contribution rate and then falls beneath it. The reform thus involves an important intergenerational aspect: it imposes an additional burden on cohorts active in the labor market near the reform but reduces the burden on later cohorts.

Figure 2. Total Contribution Rate in Reformed Pension System

Source: Authors' calculations.

Calculating the Deadweight Loss

To examine the social welfare implications of the reform, we assume that the contribution rate is viewed as a pure tax by workers and examine the deadweight losses associated with the changing contribution rates. Several points are worth noting about this approach. First, the deadweight losses associated with a given contribution rate are more complicated than we suggest here. In particular, we assume that the full contribution rate, rather than the marginal lifetime tax rate (the contribution rate net of the present value of additional benefits), is viewed as a tax. Second, we abstract from the level of redistribution embodied in the system by examining the aggregate contribution rate, rather than the contribution rates for different workers. Third, we assume that the form of the pension system (defined benefit or defined contribution, funded or unfunded) does not affect the deadweight loss associated with any given aggregate contribution rate. Finally, the deadweight losses associated with different contribution rates are not the only channel through which pension reform could affect social welfare.

Our approach thus adopts a simplified version of the deadweight loss merely as a proxy for, or as one example of, the social welfare implications of prefunding. Other approaches to modeling the social welfare implications, as long as they involve up-front costs and long-run benefits, would produce qualitatively similar results. In computing the deadweight loss we adopt the Harberger (1964) approximation, as modified by Browning (1987):

$$(21) \quad DW(Y) = 0.5\epsilon t(Y)^2 \left[\frac{w(Y)L(Y)}{1-t(Y)} \right]$$

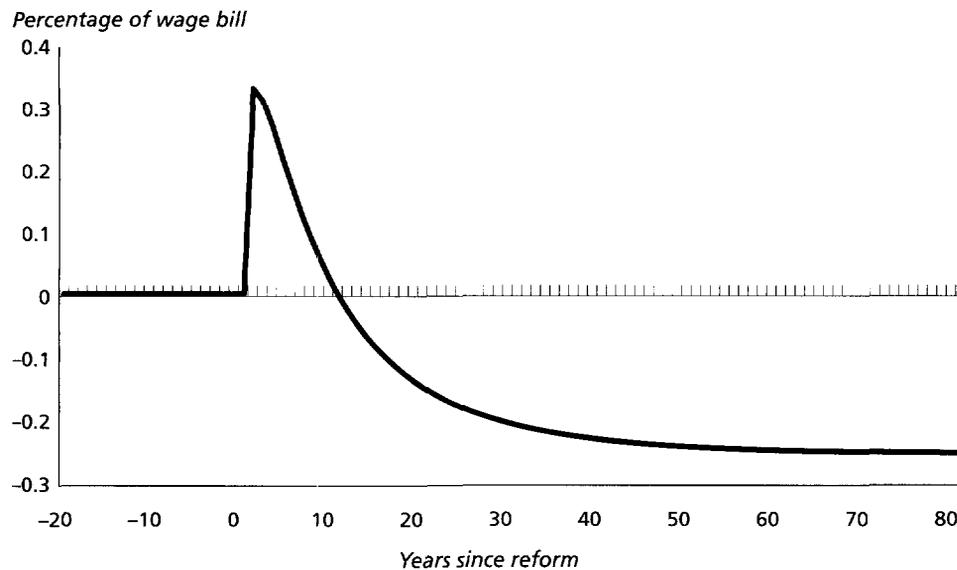
where DW is the deadweight loss Y years after the reform, ϵ is the compensated elasticity of total labor supply with respect to the net-of-tax return $(1-t)$, $t(Y)$ is the total contribution rate Y years after the reform, and $w(Y)L(Y)$ is the total wage bill Y years after the reform. Following Feldstein and Samwick (1998), we assume that $\epsilon = 0.5$. Using the same parameters as in figure 2, we obtain a time path for the deadweight loss (figure 3).

The net present value of the deadweight losses presented in figure 3 can then be compared with the net present value of the deadweight losses from continuing the pay-as-you-go system. Figure 4 presents the results given a discount rate equal to the interest rate assumed above ($r = 0.06$).

As figure 4 illustrates, the present value of the deadweight losses associated with the funded system is lower than the present value of the deadweight losses associated with the pay-as-you-go system. There is thus an expected social gain from funding. To explore the potential gain from flexible funding, we now introduce uncertainty into the model.

Parameter Uncertainty

Figure 3. Total Change in Deadweight Loss Associated with Pension Reform



Source: Authors' calculations.

To highlight the effect of unexpected shocks, we assume that at some time Z after the funding has been undertaken, a shock hits the economy that affects the death rate (d) and the productivity growth rate (g). We model two different approaches to the choices facing the policymaker: inflexible prefunding and conditional prefunding that allows the prefunding to be “undone” if warranted by the changing circumstances.

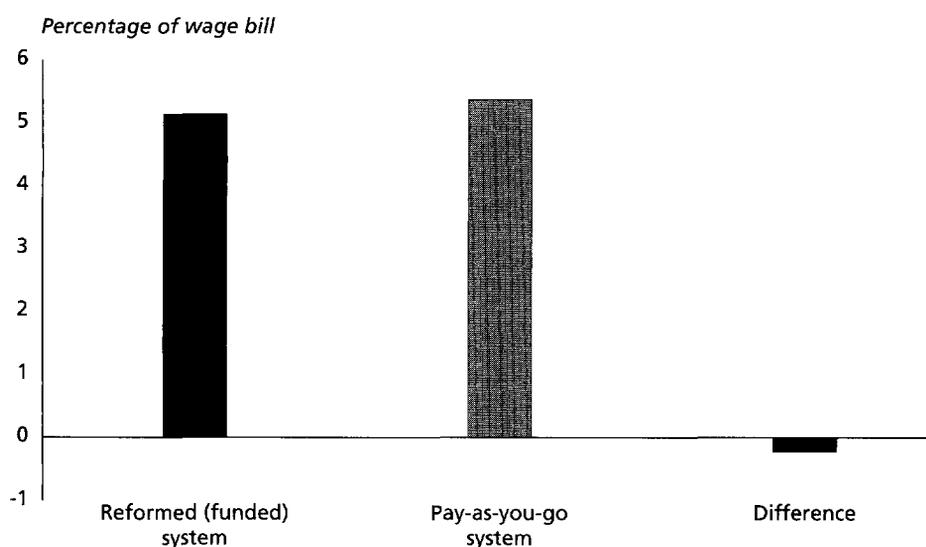
Assume that at time $Z = 30$, d rises to 0.08 and g rises to 0.035. Given these parameters, it would be worth “undoing” the funding and returning to a pay-as-you-go system. Under the conditional prefunding scenario we therefore assume that a given percentage (k) of the accumulated funding is dissipated each period by reducing the contribution rate for the funded component. Thus the conditional prefunded system evolves from an initial pay-as-you-go system to a funded system (at $Y = 0$) and then back to a pay-as-you-go system (at $Y = 30$).

More specifically, we assume that after period Z the conditional prefunding system returns to being a pay-as-you-go system. The total cost of funding the ultimate pay-as-you-go component under the conditional prefunding approach can be broken down into three components:

- Payments to pensioners who retired fewer than $Y - Z$ years ago and thus receive a pay-as-you-go pension.
- Payments to pensioners who retired between $Y - Z$ and Y years ago and thus receive a partially funded pension.
- Payments to pensioners who retired more than Y years ago and thus receive a pay-as-you-go pension.

In other words, total expenditure under the conditional prefunding system as a function of the current wage is:

Figure 4. Present Value of Deadweight Losses



Source: Authors' calculations.

(22)

$$D_t(Y, Z) = \int_0^{Y-Z} \bar{p} e^{f(t-N)} e^{-(f+d+g)s} ds + \int_{Y-Z}^Y \bar{p} e^{-\mu(Y-s)} e^{f(t-N)} e^{-(f+d+g)s} ds + \int_Y^{\infty} \bar{p} e^{f(t-N)} e^{-(f+d+g)s} ds.$$

The total pay-as-you-go contribution rate can then be calculated from equation 18. But the conditional prefunding approach can also use the prefunded reserve in reducing tax rates. We assume that the way in which the prefunded reserve is decumulated is governed by:

$$(23) \quad \tau_t(Y, Z) = \frac{D_t(Y, Z) - \kappa_t F_t(Y, Z)}{\tilde{N}(t)}$$

where F is the size of the prefunded reserve and κ is an adjustment parameter, which may vary over time to prevent the contribution rate from becoming too small. The prefunded reserve thus falls over time based on decumulations to temporarily reduce the contribution rate for the pay-as-you-go component.

We show the contribution rates for this scenario, assuming that $\kappa = 0.1$, $f = 0.015$, $N = 40$, $p = 0.4$, $r = 0.05$, $k = 0.005$, and $\mu = 0.75$ (figure 5). The contribution rate under the conditional funding approach falls sharply when the shock hits and then gradually rises. Intuitively, the higher productivity growth rate and death rate raise the attractiveness of the pay-as-you-go system (in terms of its deadweight losses) relative to a funded system.

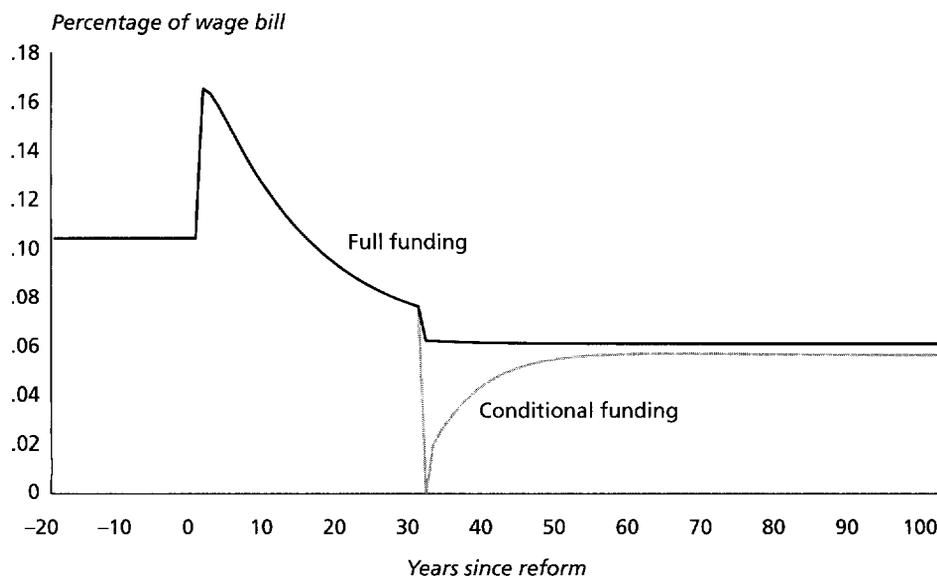
Given the new parameters (after $Y = 30$), the steady-state contribution rate under a pay-as-you-go system is lower than the steady-state contribution rate under the funded system.

The conditional funding approach takes advantage of this new environment by effectively reversing the funding, temporarily allowing a very low contribution rate. The full funding approach, which is too rigid to respond fully to the changed situation, is affected only slightly (the higher death rate, for example, reduces the contribution rate required for a given replacement rate under the funded system). The present value of the deadweight losses under the two approaches is greater under the fully funded system (figure 6).

The more formal model presented in this section delivers the same basic message as the intuitive approach presented earlier: if the future is highly uncertain, it may make sense to build in the flexibility to adjust the level of funding as the economy evolves. Several points about the model and its implications are worth highlighting.

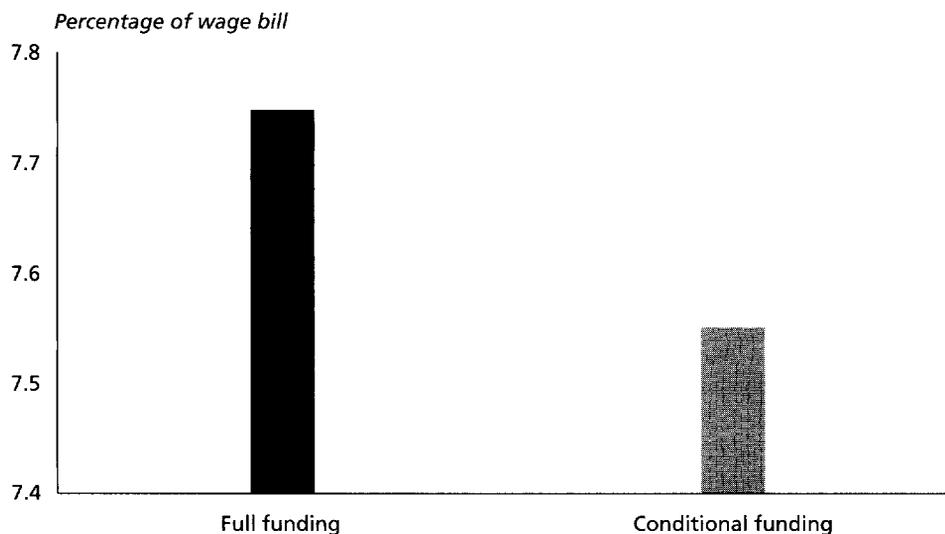
First, the main point—the desirability of flexibility in funding—shares certain similarities with the broader discussion of intergenerational risk sharing. That discussion implicitly involves a change in the level of funding, at least temporarily, when one generation is able to “share risk” with another (by reducing its funding share when a negative financial market shock hits the economy, for example). But the approach delineated here makes a somewhat different point. The discussion on intergenera-

Figure 5. Contribution Rates under Full and Conditional Funding



Source: Authors' calculations.

Figure 6. Present Value of Deadweight Losses under Full and Conditional Funding, Given a Shock



Source: Authors' calculations.

tional risk sharing focuses primarily on financial market risk (Bohn 1998 is an exception) and is concerned mostly with high-frequency fluctuations. Our approach reflects both demographic and financial market shocks, focusing primarily on low-frequency fluctuations.

Second, our model assumes that the shock hitting the economy after the reform is undertaken is permanent and large enough to reverse the relative social welfare benefits of a pay-as-you-go and funded system.⁴ On the permanency of the shock, there is some evidence that mortality rates contain a unit root (so that any given movement in mortality rates should be expected to persist until another shock hits), as assumed by Lee and Carter (1992a) and Lee and Tuljapurkar (1998). With less persistent shocks, the benefits of flexibility would be attenuated. On the size of the potential shocks, the variance in future pay-as-you-go contribution rates highlighted by Lee and Skinner (1999) and by others suggests that the uncertainty is substantial enough to warrant concern about the ability of policymakers to respond. The resolution of all uncertainty is a more restrictive assumption than is necessary to generate some benefit from flexibility—all that is required is that some of the uncertainty dissipate over time.

Third, our model examines a situation in which there is only one major reform, one permanent shock that hits the economy thereafter, and one opportunity to undo the reform. These restrictions capture some elements of reality: many countries are considering substantial pension reform over the next decade or so (and some have already undertaken reform), and significant uncertainty remains about the demographics associated with a single event—the life expectancy of the relatively large baby boom generation in many countries. But the restrictions also abstract from a more realistic setting in which several shocks hit the economy and reforms are partial rather than absolute. The restrictions are helpful in identifying the underlying issues relating to funding flexibility, but the model should not be taken too literally. The model could clearly be extended to incorporate partial funding and multiple shocks.

Fourth, we have abstracted from political concerns about different approaches to funding. Examples of the concerns often cited include corporate governance problems, political malfeasance, and dynamic consistency. Such potential concerns may be significant enough to undermine the potential benefits of a more flexible approach. If the flexibility is more likely to be abused than used sensibly, its benefits are clearly undermined.

Political Economy of Public and Private Funding

The models above suggest that funding through a flexible system may be preferable to funding through an inflexible system. If further research supports the empirical relevance of this argument, a key policy issue will be which types of pension reform are the most flexible. This section briefly explores the flexibility embedded in a public approach to prefunding relative to that in a private one.

It may be easier to achieve some degree of reversibility in a public approach to prefunding—precisely the point made by many advocates of a private approach. If adopting a mandatory, privately managed second pillar is as dramatic and discrete a step as many descriptions of that approach suggest, the substantial uncertainty surrounding its benefits should give policymakers pause.

In a partially funded system, however, it is important to have the flexibility to both increase and decrease the level of funding. The political economy of an increase in funding, given a partially prefunded system, seems unclear. Concerns about government interference in the economy are likely to increase with the size of prefunding undertaken through a public system, making it increasingly difficult to raise the level of prefunding. But increasing the prefunding in a private system may also be increasingly difficult: increases in mandatory contributions to private accounts may become politically unacceptable.

These political economy issues are important, but they go beyond the scope of this article. We note only that the degree of flexibility may need to be added as a factor—in addition to potential differences in administrative costs, risk sharing given the level of funding, and redistribution differences—in the analysis of the relative costs and benefits of public and private approaches to prefunding. A public approach that provides slightly less prefunding but more reversibility may even be preferable to a private approach that provides more prefunding but less reversibility.

A related question is how flexible prefunding can be achieved in a real-world pension system. Finland provides one example. Its system of mandatory, employer-based private pension schemes covers virtually the entire private sector workforce. These employer-based pension schemes are of several different types, governed by different legislation. The degree of funding changes as different components of the system are adjusted.

The Finnish pension system is partially funded, with some schemes—such as those for the self-employed—unfunded. (About three-quarters of all current benefits are financed on a pay-as-you-go basis.) Interestingly, Finland created a special buffer reserve, amounting to 1.1 percent of the wage bill, when it joined the European Monetary Union in 1999. The size of the buffer changes over time. The buffer is explicitly intended to provide a flexible source of prefunding: it can be adjusted in response to external shocks and can thus offset some of the adjustment costs that would otherwise be associated with membership in the European Monetary Union.

Partially funded notional defined contribution (NDC) schemes may provide another example of flexible funding. In an NDC system wage earners pay contributions based on a fixed contribution rate. The value of these contributions is credited to notional accounts, giving the system a defined contribution feature. The account values are indexed annually to some macroeconomic metric (a nominal per capita wage index in Sweden, GDP in Italy). Upon the employee's retirement, the NDC benefit is calculated by dividing the value of the account by a factor based on life expectancy.

An NDC system need not involve any broad prefunding; it could be undertaken purely on a pay-as-you-go basis. But it could incorporate some prefunding, with the degree of funding similar to a reserve requirement on accounts that can be adjusted in line with economic conditions, much as central banks adjust reserve requirements for banks. Contrary to the conclusions of Disney (1999), the flexibility of a partially funded NDC approach in coping with shocks might be a reason to prefer such an approach to other pension reform strategies.

Conclusion

In exploring the implications of demographic and economic uncertainty for pre-funding pension systems, we show that changing circumstances may warrant a change in the degree of funding. Committing to the up-front costs associated with moving from a pay-as-you-go to a prefunded system—without any assurance that the costs can be recaptured if the reform turns out to be less beneficial than had been expected—may therefore be unwise.

The model we present is intentionally limited in order to highlight the benefits of reversibility in funding. Despite its simplicity, it is consistent with sound pension policymaking and common sense.

Key areas for future research include building more detailed models in which flexibility matters, examining the empirical relevance of flexibility by parameterizing the models to real-world data, and evaluating the political economy of public and private approaches to prefunding and the implications for flexibility. This article represents a small step—but a step in a potentially important new direction that has not, to our knowledge, been explored in the pension reform literature.

Discussions of appropriate responses to the old age crisis must focus not just on expected events but also on the uncertainty surrounding those events. In the face of significant uncertainty, policymakers must carefully examine the costs and benefits of setting up inflexible funding systems. And in evaluating the costs and benefits of flexibility, they must assess whether the flexibility will be exercised judiciously.

Notes

1. Feldstein (1998) argues that such distortions already amount to approximately 1 percent of GDP for the U.S. Social Security system. As the payroll tax increases, the distortion would increase more than proportionately. The labor market distortions caused by payroll taxes within a redistributive pension system, however, are not as clear as they may initially appear (see Diamond 1998 and Orszag and others 1999).

2. According to Feldstein and Samwick (1998), “Reduction in the payroll tax rate results in a reduction in the deadweight loss that is itself equal to about 2 percent of payroll. Thus, the long-run gain from shifting to a funded system is almost as large as the entire 12 percent payroll tax. This is equivalent to a permanent increase in real income of about 5 percent of GDP” (p. 216).

3. Bohn (1998) uses “conditional prefunding” in a somewhat related sense.

4. Our example assumes that the changes in the productivity growth rate and the death rate, along with the pension system’s response to those changes, do not affect the interest rate. This lack of feedback from the real shocks to the interest rate would be warranted, for example, if the economy were a small open one. The shock hitting the economy in this case would be large enough to make it dynamically inefficient, which may seem extreme to some readers. The shock does not need to be so large to allow the conditional prefunded system to dominate the fully funded one.

References

Aaron, Henry. 1966. “The Social Insurance Paradox.” *Canadian Journal of Economics* 32 (3): 371–74.

- Baker, Dean. 1999. "Diverting the Old Age Crisis: International Projections of Living Standards in the Twenty-First Century." Washington, D.C.: Preamble Center.
- Bohn, Henning. 1998. "Social Security and Demographic Uncertainty: The Risk-Sharing Properties of Alternative Policies." University of California at Santa Barbara, Department of Economics.
- Bravo, Jorge, and Andras Uthoff. 1999. "Transitional Fiscal Costs and Demographic Factors in Shifting from Unfunded to Funded Pensions in Latin America." United Nations, New York.
- Browning, Edgar K. 1987. "On the Marginal Welfare Cost of Taxation." *American Economic Review* 77 (1): 11–23.
- Diamond, Peter. 1998. "Privatization of Social Security and the Labor Market." Paper presented at the Massachusetts Institute of Technology Public Finance lunch, 9 February, Cambridge, Mass.
- Disney, Richard. 1999. "Notional Accounts as a Pension Reform Strategy: An Evaluation." Social Protection Discussion Paper 12/99. World Bank, Washington, D.C.
- Feldstein, Martin. 1998. "Introduction." In Martin Feldstein, ed., *Privatizing Social Security*. Chicago: University of Chicago Press.
- Feldstein, Martin, and Andrew Samwick. 1998. "The Transition Path in Privatizing Social Security." In Martin Feldstein, ed., *Privatizing Social Security*. Chicago: University of Chicago Press.
- Geanakoplos, John, Olivia S. Mitchell, and Stephen P. Zeldes. 1999. "Social Security Money's Worth." In Olivia S. Mitchell, Robert J. Myers, and Howard Young, eds., *Prospects for Social Security Reform*. Philadelphia: University of Pennsylvania Press.
- Harberger, Arnold. 1964. "Taxation, Resource Allocation, and Welfare." In John Due, ed., *The Role of Direct and Indirect Taxes in the Federal Reserve System*. Princeton, N.J.: Princeton University Press.
- Lee, Ronald, and Lawrence Carter. 1992a. "Modeling and Forecasting U.S. Mortality." *Journal of the American Statistical Association* 87 (419): 659–71.
- . 1992b. "Rejoinder." *Journal of the American Statistical Association* 87 (419): 674–75.
- Lee, Ronald, and Jonathan Skinner. 1999. "Will Aging Baby Boomers Bust the Federal Budget?" *Journal of Economic Perspectives* 13 (1): 117–40.
- Lee, Ronald, and Shripad Tuljapurkar. 1998. "Stochastic Forecasts for Social Security." In David A. Wise, ed., *Frontiers in the Economics of Aging*. Chicago: University of Chicago Press.
- Orszag, J. Michael, Peter R. Orszag, Dennis J. Snower, and Joseph E. Stiglitz. 1999. "The Impact of Individual Accounts: Piecemeal vs. Comprehensive Approaches." Paper presented at the Annual World Bank Conference on Development Economics, 29 April, Washington, D.C.
- Orszag, Peter R., and Joseph E. Stiglitz. 2001. "Rethinking Pension Reform: Ten Myths about Social Security Systems." In Robert Holzmann and Joseph Stiglitz, eds., *New Ideas About Old Age Security: Toward Sustainable Pension Systems in the 21st Century*. Washington: World Bank.
- Pianin, Eric. 2000. "Drive to Fix Social Security Loses Steam; Some Economists Express Concern About Reliance on Projected Budget Surpluses." *Washington Post*, 6 February 2000, p. A4.
- Tuljapurkar, Shripad, and Ronald Lee. 1999. "Demographic Uncertainty and the OASDI Fund." [<http://www.mvr.org>].
- U.S. Social Security Administration. 2000. *Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds*. Washington, D.C.: Government Printing Office.
- World Bank. 1994. *Averting the Old Age Crisis*. New York: Oxford University Press.

Household Saving in East Asia and Latin America: Inequality, Demographics, and All That

Orazio P. Attanasio and Miguel Székely

This article compares household savings behavior in East Asia and Latin America by analyzing data on four economies—Mexico, Peru, Taiwan (China), and Thailand. It provides the first comparisons of saving in these two regions at the microeconomic level using synthetic cohort techniques. The article does not focus only on total household saving, as is common in the literature. Instead, the population is disaggregated into education groups to determine whether differences in savings behavior are related to the distribution of income. Forecasts of future aggregate household savings rates are constructed based on demographic projections. The article provides evidence that allows testing of the relevance of the life-cycle model for explaining differences in savings behavior.

Lack of efficient credit and insurance markets makes household savings a crucial determinant of welfare in developing countries. Without savings, households have few mechanisms to smooth unexpected variations in their income. Shocks may thus leave permanent scars (by interrupting human capital accumulation at early ages, for example). Since saving is one of the few means of accumulating assets in the absence of credit and insurance markets, the capacity to save becomes critical for increasing social mobility and enhancing future income-earning possibilities. Additionally, although there is controversy over the relationship between savings and economic growth, it is generally agreed that once savings start to rise—perhaps as a result of increases in income—they enhance the potential to finance investment and lead to the creation of more opportunities in the economy.

Few studies have examined savings behavior at the microeconomic level in developing countries; with very few exceptions, comparisons have focused mainly on

Orazio P. Attanasio is professor at the Institute of Fiscal Studies, University College London. Miguel Székely is research economist in the Research Department of the Inter-American Development Bank. The authors would like to thank Ricardo Fuentes for excellent research assistance, Christina Paxson for providing the Thai data and some of the programs for reading them, Martín Valdivia and Jaime Saavedra for providing the Peruvian data, and an anonymous referee for comments.

Annual World Bank Conference on Development Economics 2000

©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

aggregate savings data. In this article we use data from household surveys to compare and try to understand savings behavior at the microeconomic level in two Latin American economies (Mexico and Peru) and two East Asian economies (Taiwan [China] and Thailand). These economies were chosen because they are among the few in each region for which household data on income and consumption are available for a series of years. The experiences of these regions have differed remarkably in many dimensions over recent decades. An analysis based on household behavior may shed some light on why they have diverged.

We contribute to the literature on household savings behavior in four ways. First, to our knowledge this is the first time that savings behavior has been compared at the microeconomic level for these regions. Analyzing household saving rather than aggregate saving is useful because without microeconomic data it is very difficult, if not impossible, to interpret aggregate savings trends and to discriminate among alternative models.

Second, rather than focusing only on total household saving, as is common in the literature, we disaggregate the population into groups by education level to determine whether differences in savings behavior are related to differences in income. By focusing on education levels rather than income levels, we focus on permanent income effects rather than cyclical fluctuations. This analysis is crucial for determining whether different sectors of the population have different motives for saving and different capacities to smooth variations in income in the face of shocks and to build up assets.

Third, we use our analysis of age profiles of household saving to forecast aggregate household savings rates based on demographic projections. We are less interested in providing an efficient forecast of future savings rates than in assessing the extent to which projected demographic changes have the potential to bridge (or intensify) the differences in household saving between these two regions. The hypothesis that changes in demographic structure can affect the process of growth and saving in fundamental ways has received renewed attention. It has been argued that recent demographic shifts in East Asia—with the age groups that produce and save more growing relative to other age groups—are among the main reasons why economic performance and savings in that region improved (see Bloom and Williamson 1999 and Behrman, Duryea, and Székely 1999a). Latin America is on the verge of experiencing similar—though somewhat smaller—demographic changes, so it is of interest to verify the extent to which these shifts will contribute to increased savings in the region.

Fourth, microeconomic data allow different theories of household saving to be tested, in particular, the relevance of the life-cycle model in its different incarnations. We provide evidence that helps assess the usefulness of the theory for explaining the facts.

The evidence we present is of interest because the two Latin American countries we study experienced economic shocks during the period under analysis that are similar to those experienced recently in East Asia. Thus our analysis may provide insights into the response of saving to shocks and the capacity of different groups of the population to smooth the effects of those shocks. These insights may help policymakers protect vulnerable groups in East Asia.

We complement our analysis of savings behavior by examining a number of household decisions that are closely related to the intertemporal allocation of resources. For example, savings are likely to be affected by demographic variables through household composition effects, so heterogeneity in this dimension is likely to be important. It is also very likely that savings behavior is linked to labor supply decisions, particularly labor force participation. Additionally, if different groups of the population, such as those with different endowments of human capital, face different life-cycle earnings profiles, they will have different incentives to save. It is therefore important to characterize the behavior of different groups, as well as changes in household composition and labor force participation over time.

An analysis of this type is not without limitations. Both micro- and macroeconomic data are affected by severe measurement problems. Not only are they measured with error, but they often measure different concepts. Differences in the definition of consumption (particularly for such items as housing and health expenditure), in the population of reference (which is typically much smaller for survey data), and in the treatment of income sources (especially for pension benefits, interest income, capital gains, and imputed rents) all prevent a direct comparison between aggregate measures of savings rates and those derived from microeconomic data sources. Moreover, in the national accounts of many developing countries national saving is not disaggregated into private and public saving, and private saving is not broken down into household and corporate saving. In the few microeconomic data sources available, data on asset ownership, pension entitlements, and the like are of very limited scope and quality. Thus matching aggregate private saving with microeconomic data is not easy. Even if we think of households as the ultimate owners of corporations and assume that they are able to “pierce the corporate veil,” aggregate private saving and microeconomic data may differ if foreign investors own some firms.

In addition to these measurement problems, there are important conceptual problems. The main one is that some expenditure items, such as durables, housing, education, and health, have important savings components, but it is difficult to establish how large those components are. We cut through these issues by making some strong assumptions and trying alternative definitions of consumption and saving. In the end, however, given the data, some problems cannot be solved and the results will thus have to be interpreted with caution.

Another issue that needs to be borne in mind is that while saving is an intrinsically dynamic phenomenon, microeconomic data typically do not follow the same individuals over time because they lack a genuine longitudinal dimension. To obviate this problem, we make extensive use of the synthetic cohort techniques pioneered by Browning, Deaton, and Irish (1985). The basic idea is to follow the average behavior of groups whose membership is assumed to be fixed over time. This procedure allows us to study the dynamic behavior of the average of the variable of interest in different years. Even this technique, however, is not free from problems. The most important are the endogeneity of household formation and dissolution and the differences in mortality and migration rates across socioeconomic

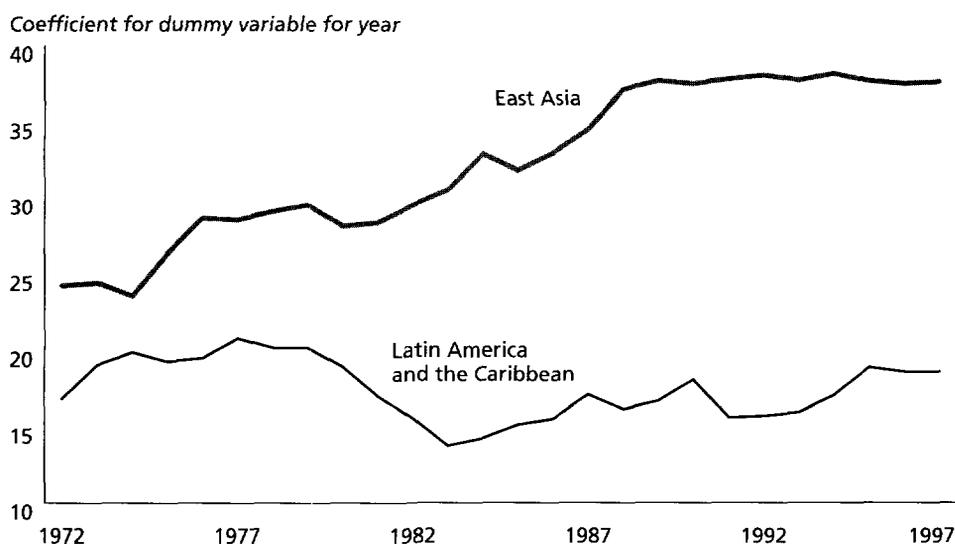
groups (see Behrman, Duryea, and Székely 1999b and Attanasio and Hoynes 2000). We discuss some of these issues below.

Aggregate Savings Trends in East Asia and Latin America

One of the main differences between the development experiences of East Asia and Latin America is that East Asia has had much higher savings rates for some time. National savings rates were already slightly higher in East Asia by 1970, and rates in the two regions have diverged dramatically since then (figure 1). (Rather than presenting means by region, which are quite “noisy,” we present smoothed profiles obtained by regressing the savings rate on a dummy variable for each year and controlling for all country characteristics.)¹ National savings rates increased in Latin America in the early 1970s, collapsed in 1982 with the onset of the debt crisis, and recovered slightly in the second half of the 1980s. In contrast, in East Asia saving increased almost continuously throughout the period, so that by 1997 the gap between the two regions’ national savings rates was about 20 percentage points.

The story is similar when we look at domestic savings as a share of GNP—and the differences even more apparent (figure 2). The average domestic savings rates for the two regions were the same in 1965, but they soon diverged and the gap has widened continuously since then. By 1997 the average domestic savings rate in Latin America was about 17 percent, while the rate in East Asia was almost 40 percent.

Figure 1. National Saving as a Share of GNP in East Asia and Latin America, 1972–97



Note: The East Asian aggregate includes only Hong Kong (China), the Republic of Korea, Singapore, and Thailand, some of the fastest growing economies with high savings rates. Taiwan (China) is not included because of lack of data. The Latin American aggregate includes all countries in the region for which data were available.

Source: World Bank 1999.

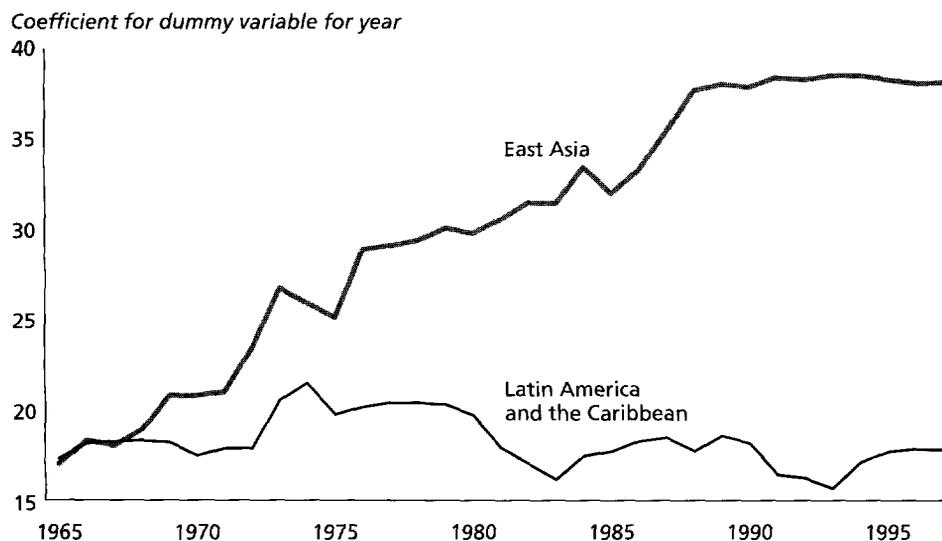
Thus most of the difference between the regions' overall savings rates can be attributed to the different patterns in domestic saving.

Several interesting features emerge from an examination of the domestic savings rates in the East Asian and Latin American economies for which data are shown in figure 2. For our purposes, the most important is that the 4 East Asian economies started out with lower domestic savings rates than the 12 Latin American countries in 1960, but they experienced a huge increase during the following 28 years. Except for Chile, no Latin American country among the 12 studied enjoyed a savings rate that was significantly higher in 1997 than in 1960. (In Ecuador and Mexico the savings rate increased, but it did so by less than 10 percentage points—a much smaller increase than in East Asia, where the average savings rate rose 30 percentage points.)

Also significant is that the savings rate in Latin America is much more volatile than that in East Asia. The four East Asian economies (Hong Kong, the Republic of Korea, Singapore, and Thailand) show a much smoother pattern and one of continuous increases. In contrast, in Latin America the savings rate changed dramatically from year to year in some countries (such as El Salvador, Nicaragua, and Peru), and in no country was there a positive trend throughout the period. By 1997 all four East Asian economies had much higher savings rates than any of the Latin American countries.

These differences are especially important in light of our microeconomic analysis, because they indicate that we are comparing two economies that have gone through intensive crisis and volatility in their savings rates with two in which domes-

Figure 2. Domestic Saving as a Share of GNP in East Asia and Latin America, 1965–97



Note: The East Asian aggregate includes only Hong Kong (China), the Republic of Korea, Singapore, and Thailand, some of the fastest growing economies with high savings rates. Taiwan (China) is not included because of lack of data. The Latin American aggregate includes all countries in the region for which data were available.

Source: World Bank 1999.

tic saving has increased continuously and smoothly. In interpreting our results, we will have to keep this in mind.

At the same time that domestic savings patterns in East Asia and Latin America were diverging, there were also significant differences in other dimensions. Total fertility rates declined much faster in East Asia than in Latin America. In 1965 the demographic structures in the two regions were almost the same, but since then there has been a growing gap. Fertility in both regions started declining in the 1950s, but the reduction was much faster and steeper in East Asia, and by 1965 the cohorts entering working age were much larger than the newborn ones (United Nations 1998). Thus since 1965 a much larger share of the East Asian than the Latin American population has been entering the ages in which savings rates rise. This composition effect may be an important force behind the differences in saving.

The reductions in fertility in the two regions are highly correlated with sharp increases in female labor force participation. Fertility and female labor force participation are usually jointly determined, and they have a double effect on savings behavior: lower fertility rates mean that households have fewer children on average, while higher labor force participation means that more household members are in the workforce (and thus income is higher). The result of both effects is higher household per capita income and thus higher savings capacity. Female labor force participation was lower in Latin America than in East Asia in 1960 (World Bank 1999). Although it has increased slightly faster there than in East Asia, the gap between the regions remains large.

Another important transformation taking place in the two regions at the time their savings patterns were diverging was the transition toward higher levels of schooling. Here too East Asia enjoyed much faster progress than Latin America. In 1960 average schooling in East Asia was 0.7 years longer than in Latin America; by 1990 the difference had grown to 2.5 years (Behrman, Duryea, and Székely 1999b). More educated individuals usually have higher incomes and thus higher savings capacity, so this is another potentially important factor behind the differences in domestic saving.

Yet another difference between the two regions has been the rate of growth. It is not clear whether economic growth precedes higher savings rates or higher savings rates precede growth. But since the mid-1970s East Asia has had higher GDP per capita (adjusted for purchasing power parity). Differences in growth rates have been apparent since the mid-1960s, precisely when domestic savings rates surged in East Asia.

Methodology and Data

The main purpose of the analysis presented in the following sections is to characterize the patterns of household saving over the life cycle. The conceptual framework is the life-cycle model, although we do not take a stance on the version of the model (with or without precautionary saving, liquidity constraints, bequest motives, or habit formation) that might be most appropriate to describe the data. In addition to the life-cycle profiles of saving for the population at large, we focus on the differences in savings behavior among different groups of households. We also relate

the observed savings behavior to other variables, particularly demographic ones, which are likely to be important determinants of saving.

This type of analysis is useful for two reasons. First, the focus on different groups of the population—characterized by different earnings profiles, demographics, and shocks over the sample period—could shed light on the determinants of saving. We also emphasize important differences across groups that remain hidden in aggregate analysis. This is particularly important for Latin America, which is characterized by substantial inequality. Second, giving age profiles of saving a semistructural interpretation allows extrapolation of the relationship that links savings rates to demographic variables—and thus forecasts of future household savings rates.

The type of microeconomic data available dictated the technique we use. Unfortunately, most Latin American countries with household surveys that provide information on both income and consumption include only one or two data points. Because we believe that cohort effects are particularly important, we focus on the two Latin American economies (Mexico and Peru) and two East Asian ones (Taiwan [China] and Thailand) for which data are available for several points in time. Because we study a dynamic phenomenon and use time series of cross-sections, we are forced to use synthetic cohort techniques. These techniques allow us to follow the same groups of individuals over time, even in the absence of a genuine longitudinal dimension in the data.

Methodology

In the presence of strong cohort effects, the age profile of savings rates, consumption, or any other variable in a single cross-section may not correspond to the age profile of any individual. To circumvent this problem, we group the observations in each of several repeated cross-sections according to one or more variables chosen so that the group membership from which the observations are drawn is likely to be fixed. In this way, rather than following the behavior of individuals over time, we follow the average (or any other moment) of the variables of interest for the groups. In the context of a life-cycle model an obvious way to form groups is on the basis of the year of birth of the household head, so that we can follow the behavior of different cohorts as they go through different phases of their life cycle. We also consider education groups, under the assumption that the accumulation of human capital can be done only in the early phases of the life cycle.

While synthetic cohorts are extremely useful, the technique is not free from problems, since group membership might change over time and household formation and dissolution could be endogenous to the phenomenon under study. Different mortality and migration rates can also induce changes in composition. Endogeneity of household formation and dissolution is relevant if the propensity to form a household at the beginning of the life cycle is different in different groups of the population and household dissolution results in elderly individuals going to live with their offspring. Extended families and family arrangements in old age are particularly relevant for our analysis because they are directly related to life-cycle saving and the incentives to save.

Most of the analysis we present here is graphical. To identify the life-cycle profile of several variables of interest, we plot the average data for each cohort against age. Because different cohorts are observed over different parts of their life cycle, we are able to track the age profiles. Moreover, if the sample period covered by the time series of the cross-section is longer than the interval used to define a cohort, we observe different cohorts at the same age (although obviously at different points in time).

An important identification caveat should be kept in mind. While it is true that with a sufficiently long sample period one observes different cohorts at the same age, one should resist the temptation to interpret the resulting differences as always due purely to cohort effects. The obvious reason is the possible presence of year effects. We use smoothing techniques to present age and cohort effects, but strictly speaking, age and cohort effects can never be disentangled without additional information or restrictions from time effects because of the exact linear relationship linking age, time, and year of birth. While in some cases, such as demographic variables, it is natural to impose the absence of year effects, in other cases this assumption is a strong one and the results should be interpreted with caution. One should always remember that any combination of cohort and age effects can be obtained as a combination of age and time or time and cohort effects. We discuss these issues below.

Once we estimate the age effects for savings rates, we extrapolate them to forecast future aggregate savings rates. In particular, we use the following relationship. If we indicate the aggregate savings and income at time t by S_t^{ag} and Y_t^{ag} and the savings, income, and size of group c (cohort) at time t by S_t^c , Y_t^c , and N_t^c , the aggregate savings rate will be given by the following expression:

$$(1) \quad s_t^{ag} = \frac{S_t^{ag}}{Y_t^{ag}} = \frac{\sum_c S_t^c N_t^c}{\sum_c Y_t^c N_t^c} = \sum_c \frac{S_t^c}{Y_t^c} w_t^c$$

where

$$w_t^c = \frac{Y_t^c N_t^c}{\sum_c Y_t^c N_t^c}$$

If we assume that group savings rates are a function of age and cohort effects, we can extrapolate from the group savings rates

$$\frac{S_t^c}{Y_t^c}$$

estimated in the microeconomic data using equation 1, relative income profiles, and demographic projections to forecast future savings rates. These forecasts should be treated with extreme caution, however, because they are based on the behavior in a given economic environment (that is, by households facing given earnings profiles and demographic profiles and a given set of institutions, including arrangements for

old age). Moreover, the data problems mentioned above make it very difficult to match the micro- and macroeconomic measures of saving. Nonetheless, these forecasts are informative about the potential effects of demographic trends and changes in the composition of the population on aggregate saving.

If the groups are defined not only by the year of birth of the household head but also by educational achievement, forecasting using the aggregate savings rate in equation 1 becomes even harder. This is because it involves forecasting not only the age structure of the population (for which demographic projections can be used) but also the accumulation of human capital. General equilibrium effects—in particular, the effect that the relative size of different skill groups might have on the returns to human capital—complicate this type of exercise even further. In principle, in the absence of cohort effects one can use the age profile from a single cross-section to perform the same exercise. Caution is needed in interpreting the results of such an exercise, however.

Even if one does not want to disentangle age and cohort effects and consider equation 1 at two points in time, one can use equation 1 to decompose the changes in aggregate savings rates into changes due to shifts in the cross-sectional age profile and changes in the weights:

$$(2) \quad s_t^{ag} - s_{t-1}^{ag} = \sum_c S_t^c (w_t^c - w_{t-1}^c) - \sum_c w_{t-1}^c (s_t^c - s_{t-1}^c).$$

Changes in weights can in turn be decomposed into changes in the relative sizes of different age groups and changes in their disposable income (an exercise we perform later).

Data and Definitions

Few high-quality data sources include information on both income and consumption; many of the sources that do so include data far apart in time. The problem affects both industrial and developing economies. For Mexico, Peru, Taiwan (China), and Thailand cross-sectional data observed at several points in time are available. For Mexico we have data from the National Survey on Income and Expenditure of Households, conducted by the National Institute of Statistics, Geography, and Informatics, for 1984, 1989, 1992, 1994, and 1996. The last year is of particular interest because it is the year after the peso crisis. The data for Peru are from the National Household Survey for the Measurement of Living Standards and cover 1985, 1991, 1994, and 1997. They include data from two surveys that surround the 1990 crisis. The data for Taiwan (China) are annual data for 1976–96 from the Family Income and Expenditure Survey, conducted by the National Statistics Office of Taiwan (China). The data for Thailand—for 1975, 1981, 1986, 1988, 1990, 1992, 1994, and 1996—are from the Socio-Economic Survey (SES) conducted by the National Statistical Office.²

Using these data, we construct measures of income, consumption, household composition, educational attainment, and labor supply. Some of the definitions relevant for the analysis of saving include the following:

- Income is defined as disposable household income. It includes earnings, transfers, capital income, and nonmonetary income.
- We use four definitions of consumption to calculate savings rates. For the first, s_1 , we include all household expenditure; this is closest to the definition typically used in national accounts data. The second savings rate, s_2 , excludes from consumption all expenditures on durable goods as well as expenditures on health and education. This measure tries to take into account the fact that some expenditure items have an important savings component. While this definition is far from perfect—it does not, for example, include services that accrue from durable goods—the analysis of consumption and saving based on this definition deserves attention. The third definition, s_3 , excludes expenditures on durable goods but considers health and education current consumption rather than savings, as in s_2 . Finally, a measure we label s_4 includes durable and nondurable expenditures in the definition of consumption but excludes health and education, which are considered savings.
- All surveys include some definition of human capital. We divide the population into three groups: primary education or less, some secondary schooling, and higher education. This classification takes into account institutional differences across economies.
- Household arrangements differ across economies. While we present some evidence on this and document the extent of possible problems with endogenous household formation and dissolution, in the end we use the standard definition of declared household head across economies.

Throughout the analysis we divide the samples into year-of-birth cohorts. To work with cells of reasonable size, we use a five-year definition. The cohort definition is homogeneous across economies.

Cell sizes range from 86 to 2,528 households. Interestingly, one of the smallest samples is observed in Mexico, the country with the largest population: Mexico has about 70 million more people than Taiwan (China), yet we have about 50 percent more observations for Taiwan (China). The differences are larger at older ages. Since Mexico and Peru have the smallest samples, estimates derived for these economies will have less precision, especially for the older cohorts, than those for the two East Asian ones. Except for Thailand (where weights are not available), we use population weights to compute our results.

Static Analysis of the Most Recent Data

The data on aggregate private savings rates presented in figure 2 are far from perfect, for many reasons. One of the main problems is that in the national accounts savings are calculated as residuals of other aggregates that are also measured with some error. In this section we first complement the aggregate evidence by presenting household savings rates calculated from the microeconomic data for the most recent year available for each of the four economies. We also characterize

saving for different population groups and investigate whether the differences in savings rates between East Asia and Latin America are due to demographic differences.

To link the aggregates presented earlier and the results from the microeconomic data, we concentrate on the most standard definition of savings—total disposable income minus total expenditures, divided by total disposable income (s_1). The figures we present are computed as ratios of average saving and average income, rather than the average of the ratio (table 1). This procedure reduces the effect of outliers.

Based on the aggregate data, Thailand appears to have the highest domestic savings rate (almost 36 percent), while Mexico, Peru, and Taiwan (China) show roughly similar savings rates. When we turn to the microeconomic data, a different picture emerges. Mexico and Peru register an s_1 of about 9.5 percent, while Thailand and Taiwan (China) have rates of 29.7 percent and 49.1 percent. Thus Thailand has a much higher rate of domestic saving than Mexico and Peru, and much of the difference seems to originate at the household level; the difference between Taiwan (China) and the Latin American countries is even greater. So, if we set aside the issue of the comparability of micro- and macroeconomic sources, the small difference at the aggregate level must reflect lower public and corporate saving in Taiwan (China).

As expected, dependency ratios are much higher in Mexico and Peru (by about 20–30 percentage points), reflecting the fact that these countries are at an earlier stage of the demographic transition, with a smaller proportion of the population of working (and saving) age. Thailand has a much higher female labor force participation rate than Mexico and Peru; since higher participation is associated with higher savings capacity, the result is consistent with the huge difference in household savings rates. Thailand also registers fewer average years of schooling than the two Latin American countries and similar per capita GDP; these variables therefore cannot account for the difference. Institutional factors, such as the lack of compulsory retirement benefits in Thailand until 1999 and the tradition of public pension provision in Mexico and Peru, may partly explain the difference in household savings rates.

Table 1. Household Saving and Selected Other Indicators, by Economy, 1996

(percent, except where otherwise indicated)

<i>Economy</i>	<i>Aggregate domestic savings rate</i>	<i>Household savings rate</i>	<i>Youth dependency ratio</i>	<i>Female labor force participation rate</i>	<i>Average years of schooling</i>	<i>PPP GDP per capita (international dollars)^a</i>
Mexico	25.4	9.5	0.59	44.4	7.1	5,757
Peru ^b	19.0	9.6	0.60	64.5	8.5	2,993
Taiwan	26.8	49.1	0.31	57.4	9.3	14,634
Thailand	35.9	29.7	0.41	79.7	5.8	5,080

Note: Data on aggregate domestic saving and PPP GDP per capita for Mexico and Peru are from World Bank (1999).

a. Gross domestic product converted to international dollars using purchasing power parity (PPP) rates.

b. Data for Peru refer to 1997.

Source: Authors' calculations based on household survey data. Data for Taiwan (China) are from the National Statistics Office of Taiwan (China). For Taiwan (China) data on aggregate domestic saving refer to gross national savings, and GDP per capita is not adjusted for purchasing power parity.

For Taiwan (China), the lower dependency ratio, the greater average years of schooling, the larger per capita GDP, and the higher female labor force participation rate (relative to Mexico, although not to Peru) are consistent with the huge difference in household savings rates between Taiwan (China) and the two Latin American economies.

Thus one possible explanation for the differences in household savings rates is that the East Asian economies are at a later stage of the demographic transition, with larger shares of their populations at ages at which savings rates typically peak. For all ages the savings rate is higher in Taiwan (China) and Thailand than in Mexico and Peru. The weight of the middle-aged groups is also somewhat greater, as expected. As a gross approximation to assess the role of demographics in the differences in saving, we recompute the savings rates in Mexico and Peru using weights from Taiwan (China) and vice versa. Although the rates for Mexico and Peru increase and the rate for Taiwan (China) declines, the difference relative to the original savings rate in each economy is only marginal. This suggests that having a larger share of household heads in age groups that save more does not account for the major differences in saving.

This exercise only partially accounts for demographic differences, because it considers the distribution of households across working ages but ignores the fact that the major source of demographic differences is the larger share of the population in Latin America in the 0–15 age range. Both Latin American countries have smaller population shares of prime working age than the East Asian economies. A second accounting exercise attempts to take this into account, at least in a very general way. The calculations consist of weighting the age-specific savings rates in Mexico, Peru, and Thailand with the total population weight of Thailand and weighting the Taiwan (China) rates with the Peruvian weights. Doing so narrows the difference between the Latin American countries and Taiwan (China) from 37 percentage points to about 14 percentage points, which gives us a gross idea of the importance of differences in demographic structure.

In examining savings rates by education group, we observe that Taiwan (China) and Thailand have much higher rates than Mexico and Peru, with the largest differences between Thailand and the Latin American countries among the most educated households (Thai savings rates are 28 percentage points higher for the most educated households, 22 percentage points higher for the least educated). Clearly, the differences in household savings rates are given not by the size of the education groups but by the differences in the group-specific rates. The simple accounting exercise in which we recompute the Latin American averages using weights from Taiwan (China) and vice versa confirms this.

In comparisons of saving by income quintile, the most striking difference between Thailand and the Latin American countries is at the top of the distribution, where savings rates in Thailand are considerably higher (table 2). Differences appear beginning in the second quintile, but the difference is greatest in the top quintile. Surprisingly, the only group in which Mexico has higher (less negative) savings rates than Thailand is the poorest quintile. This may indicate the importance of transitory

income components in the two countries or larger measurement error for the lowest income quintile. Much of the difference between the aggregate savings rates in these two countries is explained by the extremely high savings rates among Thai households in the two richest quintiles.

The only economy in which differences along the income distribution are small is Taiwan (China), where there is only a 15 percent gap between the savings rates of the poorest and richest quintiles. This is not surprising, since inequality in Taiwan (China) is very low (as indicated by the Gini coefficient of just 0.30).

In comparisons of the quintile savings rates in Taiwan (China) with those in each of the other economies, the most interesting feature is that in all cases the largest difference is in the poorest quintile. The comparison between Peru and Taiwan (China) is especially illustrative. Almost all of the difference in the total household savings rate (about 40 percentage points) is explained by the disparities in the first four quintiles. In fact, the richest 20 percent of households in Peru have a savings rate only 15 percentage points lower than that of their counterparts in Taiwan (China). Although the differences between Mexico and Taiwan (China) are less extreme, the conclusion still holds that the largest part of the difference in household savings comes from poorer households.

Differences in the relative importance of education groups do not, then, account for the large disparities in household saving between East Asia and Latin America, while differences in demographic structure have the potential to explain part of the gap. But an important part of the story is the difference in savings rates across the income distribution. It is difficult to determine which way the causality runs (do all groups of society have similar earnings and savings potential when income is more equally distributed, or do income-earning assets increase when more individuals are able to save for reasons other than income?). But regardless of the direction of causality, the finding is important because it reveals that in Mexico, Peru, and Thailand the richest sectors of the population have much greater capacity to accumulate assets and smooth unexpected variations in income than the rest of the population, while in Taiwan (China) this is not the case. In Mexico and Peru nearly all household savings originate among the richest 20 percent of the population. After

Table 2. Household Savings Rate by Income Quintile, 1996

<i>Economy</i>	<i>Total</i>	<i>1</i> <i>(poorest)</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i> <i>(richest)</i>	<i>Gini</i> <i>coefficient</i>
Mexico	9.5	-0.17	-0.05	-0.01	0.05	0.21	0.53
Peru ^a	9.6	-1.81	-0.52	-0.18	0.07	0.40	0.51
Taiwan (China)	49.1	0.39	0.43	0.46	0.48	0.54	0.30
Thailand	29.7	-1.32	0.00	0.15	0.29	0.49	0.53
Taiwan (China)-Mexico difference		0.55	0.48	0.46	0.43	0.33	
Taiwan (China)-Peru difference		2.20	0.95	0.63	0.41	0.15	
Taiwan (China)-Thailand difference		0.71	0.43	0.30	0.19	0.05	

a. Data are for 1997.

Source: Authors' calculations based on household survey data.

accounting for these differences along the distribution, it is much less surprising that Taiwan (China) has savings rates well above those registered in the two Latin American countries.

Aggregate Household Saving over Time

Savings rates changed significantly over the 28-year period studied. In Mexico s_1 increased between 1984 and 1989, a period characterized by stagnation and the partial recovery of growth after a substantial decline in 1986. The 1989–94 period was characterized by a consumption boom and a simultaneous increase of 3 percentage points in the savings rate. Changes in s_2 , s_3 , and s_4 were similar in 1984–89, but in 1989–94 the increase was much smaller than that in s_1 and there was even a decline in some cases. This pattern suggests that the consumption of durable goods was increasing faster than expenditures on nondurables, education, and health. Between 1994 and 1996 average s_1 and s_3 declined sharply, while average s_2 and s_4 rose. The medians for all four definitions of savings rates declined during this period. They increased much less than the average or declined between 1984 and 1989 and dropped in 1992. The savings rates for the poorest 50 percent of the population were less responsive to the increase in income after 1989, while the 1995 tequila crisis had a larger negative effect on their savings, perhaps because of the limited income-smoothing mechanisms available to this group.³

Household saving in Peru has been much more erratic. Average rates increased between 1985 and 1991, declined in 1994, and rose in 1997. This trend is surprising, since Peru experienced sharp declines in per capita GDP during the 1980s and more stable growth during the first half of the 1990s. The only period in which savings rates behaved as expected was 1994–97, when per capita GDP grew at a much faster pace. In the case of median rates, the data reveal that the 1990 crisis had a larger negative effect on the poorest 50 percent of the population, while the 1994–97 recovery did not have the positive effect on this group that was observed at the upper part of the income distribution.

The picture for Taiwan (China) and Thailand is quite different. Comparing 1985–86 and 1996, we find that with few exceptions, household saving increased smoothly. The four savings rates peaked in 1996 at 49.1 percent, 84.0 percent, 74.8 percent, and 58.5 percent in Taiwan (China) and at 33 percent, 62.3 percent, 56.9 percent, and 38.8 percent in Thailand. The patterns followed by the average savings rates and the median rates show that the increases occurred across definitions and across the income distribution.

We use annualized household savings from the surveys to estimate the relative importance of household savings at the national level. These estimates are only gross approximations: household surveys normally suffer from misreporting or underreporting of income, so the data may not reflect total household savings with precision. Furthermore, the calculation does not account for the savings that households have in firms, the value of pensions, or other important items. The results are nevertheless useful for identifying differences across countries.

As expected, household savings account for the largest share of GDP in Taiwan (China). Interestingly, however, although household savings in Mexico are lower than those in Thailand and similar to those in Peru, they represent a much larger share of GDP. The results confirm that a very large part of total domestic savings in Taiwan (China) originates at the household level.

We use equation 2 to decompose the changes documented earlier into three effects (table 3). The first is an age profile effect, which reflects the increase in saving due to the fact that cohorts age through their life cycle and save more or less depending on their needs and future prospects. The second is a demographic effect, which measures the change in saving due to the fact that the population weight of different age groups shifts. The third is an income effect, which reflects the fact that as individuals age, their income tends to rise and they thus have greater savings capacity.

The contrast between East Asia and Latin America is stark. In Mexico and Peru all of the change in household saving is driven by the age profile effect of cohorts moving through the life cycle; in Taiwan (China) and Thailand the age profile effect is positive, but most of the change in saving is accounted for by the demographic and income effects. These results support the view that the East Asian economies have experienced much larger increases in saving because of the increase in their income and the shift in the population toward age groups that save more, not because of households moving through the life cycle.

It must be kept in mind that the time span in table 3 is longer for East Asia than for Latin America and, more important, that the period under analysis in Latin America is characterized by economic instability and low economic growth. A large part of the difference may thus reflect the fact that the economic environment in East Asia has been much more favorable for building up savings.

A Life-Cycle Analysis of Household Savings Behavior

We are now ready to analyze the life-cycle patterns of household saving in the four economies under study. To do so, we use the time series of repeated cross-sections to construct synthetic panel data. We divide each survey into 12 birth cohorts and 3 education groups. Much of the analysis in this section is graphical. In particular, we plot the life-cycle profiles of several variables of interest. It is therefore worthwhile to briefly describe the way in which we construct the graphs.

Table 3. Decomposition of the Change in Household Saving

(percentage points)

Economy	Period	Total change	Age profile effect	Effect of change in weights	Effect of change in weights	
					Demographic	Income
Mexico	1984–96	1.0	1.3	–0.2	–1.7	–1.5
Peru	1985–97	12.7	14.3	–1.7	–1.8	0.1
Taiwan (China)	1976–96	20.0	3.74	16.3	7.6	8.6
Thailand	1975–96	19.5	9.03	10.4	4.6	5.8

Source: Authors' calculations based on household survey data.

We plot the cohort averages at different points in time against the age of the household head and connect the points for each cohort. Visually, we follow the average behavior of each cohort as it ages. If the interval that defines a cohort is shorter than the period covered by our sample, we observe different cohorts at the same age. While it is tempting to interpret such a difference as a cohort effect, one should remember that these figures refer to different years and therefore could be equally explained by a time effect.

Changes in Household Formation and Composition

We start the analysis by looking at household formation. For this reason the first graph we present is based on individual data. In each panel of figure 3 we plot the average age of the head of the household in which an individual lives against the age of the individual. To produce this graph we use all individuals in the sample, regardless of their position in the household. For each economy we display four panels: the top-left panel graphs the data for the entire sample, while the other three graph the three education groups.

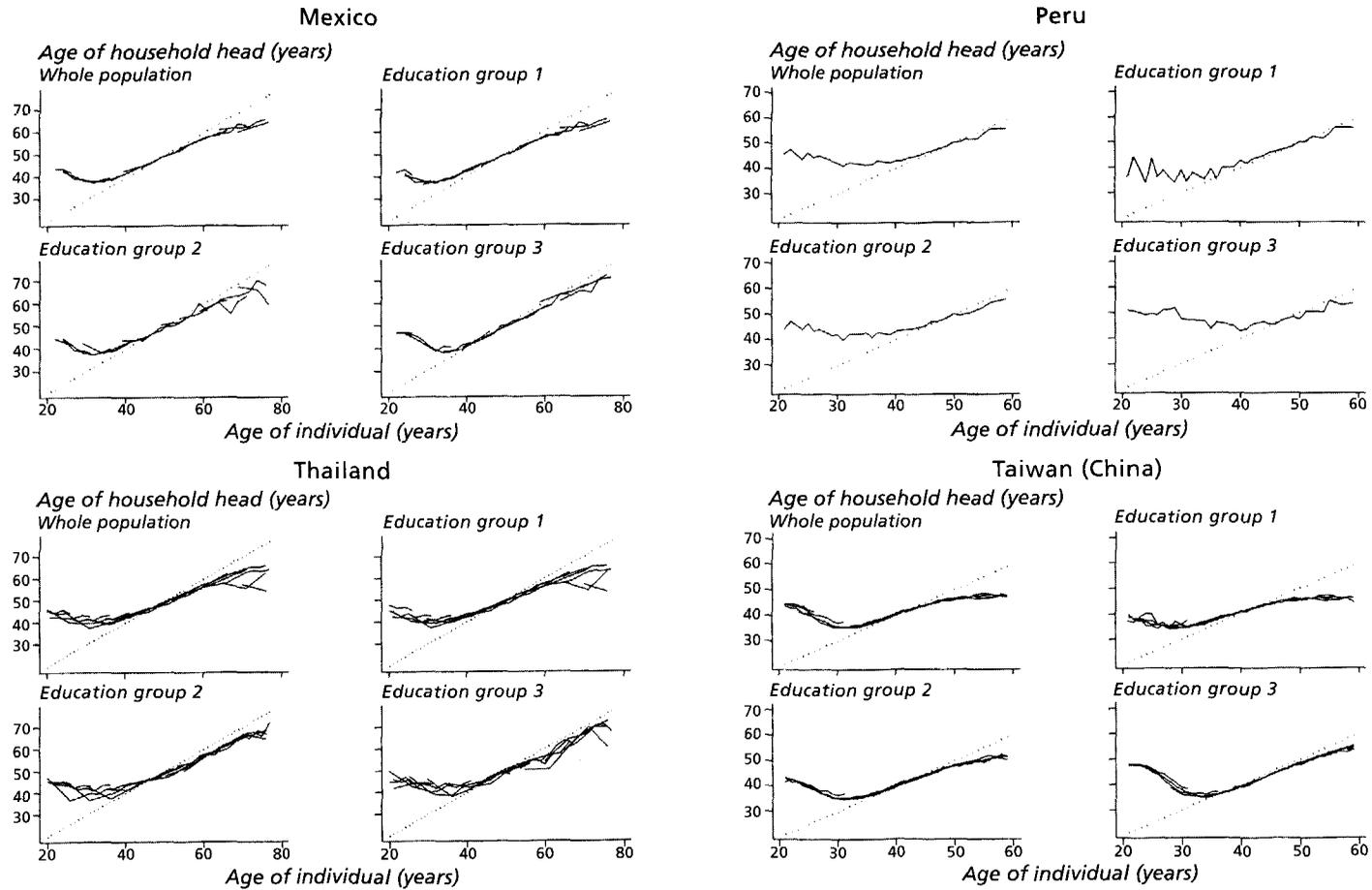
If all individuals in a given cohort were household heads (or living in a household with a head of the same age), the 45-degree line would coincide with the cohort profile of the average household age. In figure 3 the cohort profiles diverge from the 45-degree line at the beginning and the end of the life cycle. Naturally, headship rates are low at the beginning of the life cycle, so that the cohort profile lies above the 45-degree line, indicating that some young adults still live with their parents. The speed with which the profile approaches the 45-degree line is an indication of how early new families are formed. Toward the last stages of the life cycle headship rates decline and fall below the 45-degree line, because the elderly merge into households in which the head is younger.

Two patterns are apparent. First, in all four economies there are strong differences across education groups, especially in the extent to which the profile falls below the 45-degree line. Less educated individuals seem to be more likely to move in with younger household heads. In contrast, the most educated have much higher headship rates and continue to be heads of households even at older ages. These findings suggest that the family plays an important role in smoothing consumption for the elderly, especially among those with lower income-earning capacity.

Second, there are strong differences across economies at both ends of the life cycle. In particular, the phenomenon of elderly individuals living in households headed by younger ones is much more prevalent in the two Asian economies. In contrast, at the beginning of the life cycle the differences cut across regions. Household formation seems to occur much later in Peru than elsewhere. At the opposite extreme is Taiwan (China), where by age 30 the profile already coincides with the 45-degree line.

The results shown in figure 3 are also interesting from a methodological standpoint, because they show that household composition changes in important ways through the life cycle, especially among the least educated. This means that even though we are tracing the same type of household in the repeated cross-sections, the

Figure 3. Age of Household Head for Household in Which Each Individual Lives



Source: Authors' calculations based on household survey data.

composition of the group is changing, blurring our inferences about the behavior of cohorts as they age.

Figure 4 plots the average years of schooling of household heads, as well as the proportion of household heads with secondary and higher education. Since after age 26 very few individuals continue to acquire formal education, we plot the cohorts from age 26 on. If there were no composition effects in the cross-sections under analysis, there would be differences in levels of education across cohorts, but the age pattern of each cohort would be close to a horizontal line.

Two patterns emerge from figure 4. First, the cohort effects for average years of schooling and for the proportion of individuals with secondary and higher education seem to be much larger in East Asia. Second, there are some compositional changes in the surveys. While some of these changes could be attributed to sampling error, systematic positive trends in years of education or the proportion of well-educated individuals could be a symptom of differential mortality effects. While these effects exist for Mexico, Taiwan (China), and Thailand, they are not particularly strong or significant. There are significant shifts in Peru, however, suggesting that the results for this country should be interpreted with more caution.

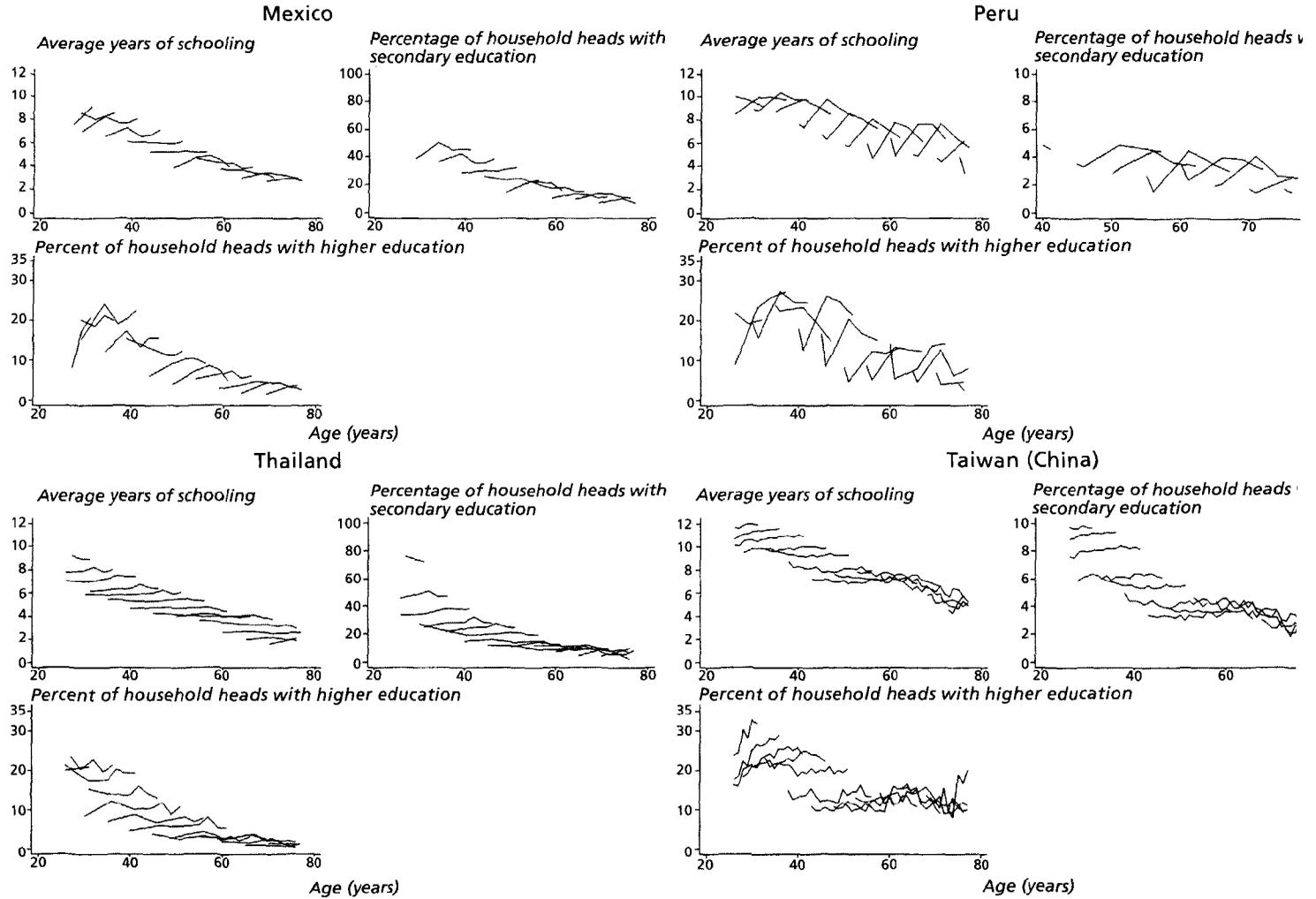
In figure 5 we analyze household data on household size. In the four panels for each economy we plot the log of household size against the age of the household head (for the entire population and the three education groups). Three features deserve emphasis. First, there are large differences in household size among education groups, with the least educated having the largest families in all cases. Second, there are large cohort effects, especially for Taiwan (China) and Thailand, with the youngest cohorts having much smaller families. Third, families are much smaller in the East Asian economies. In Peru and Thailand household size declines more slowly in the last part of the life cycle than in Mexico and Taiwan (China), perhaps because children leave home much later or because older adults join what become extended families.

The patterns in figure 5 are mirrored in figure 6, which plots the average number of children against the age of the household head. It is not implausible to interpret the differences between cohorts shown in figures 5 and 6 as pure cohort effects, since it is plausible to rule out the existence of systematic year effects.⁴ Cohort effects are stronger in the two East Asian economies, and the number of children is smaller. The differences across economies tend to be larger in the groups with primary and secondary schooling.

Income, Consumption, and Savings Profiles

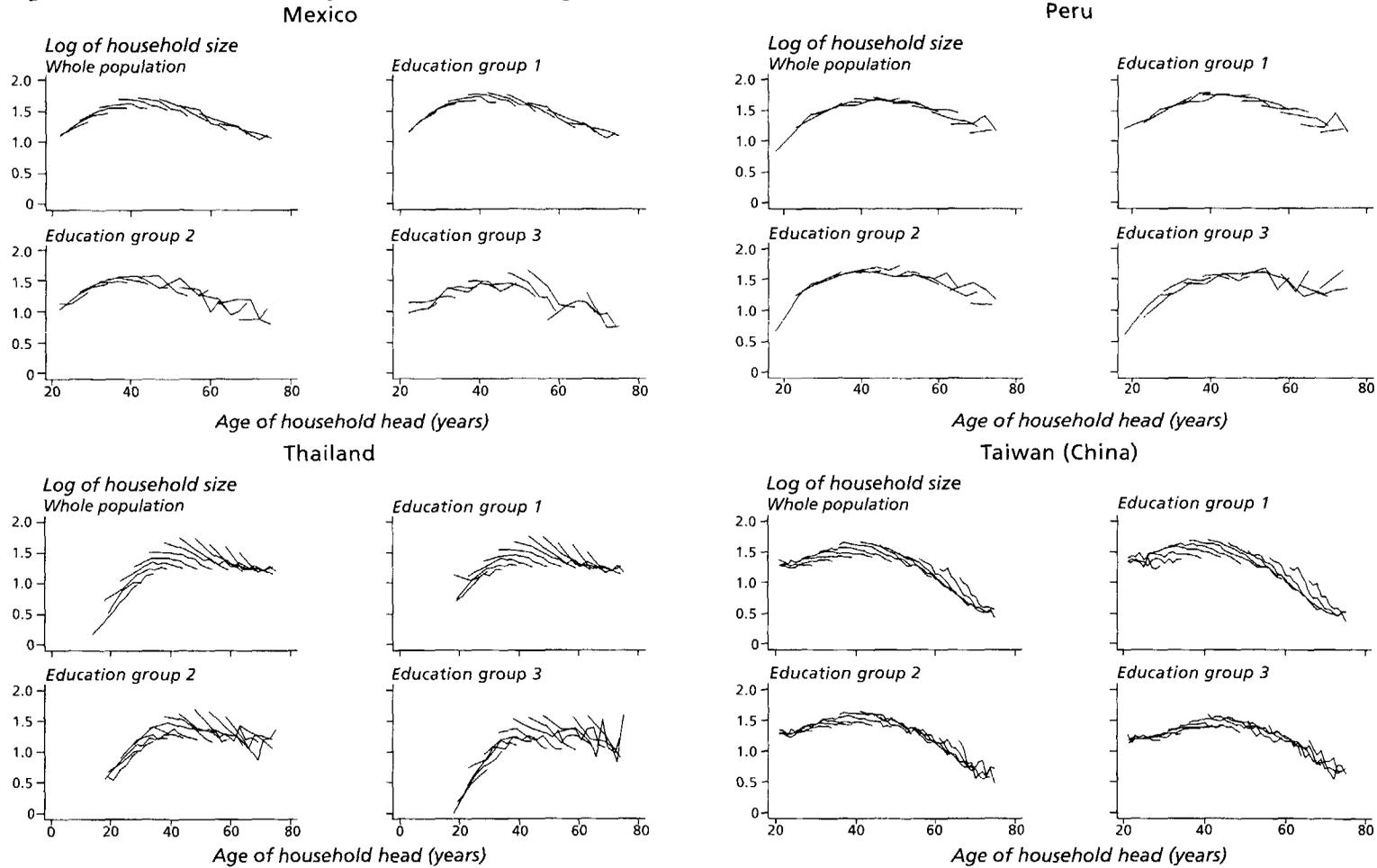
In figure 7 we plot the average of the log of disposable household income and the log of total consumption expenditure (note that the scale is different for each economy). In all four economies, as in other countries, consumption and income track each other closely. Moreover, differences in the shape of the income profile across education groups are mirrored in differences in the consumption profile.⁵ Once again there are important differences across both education groups and economies. First, the most educated not only have higher incomes but their income profile is

Figure 4. Average Years of Schooling of Household Heads and Share with Secondary and Higher Education, by Age



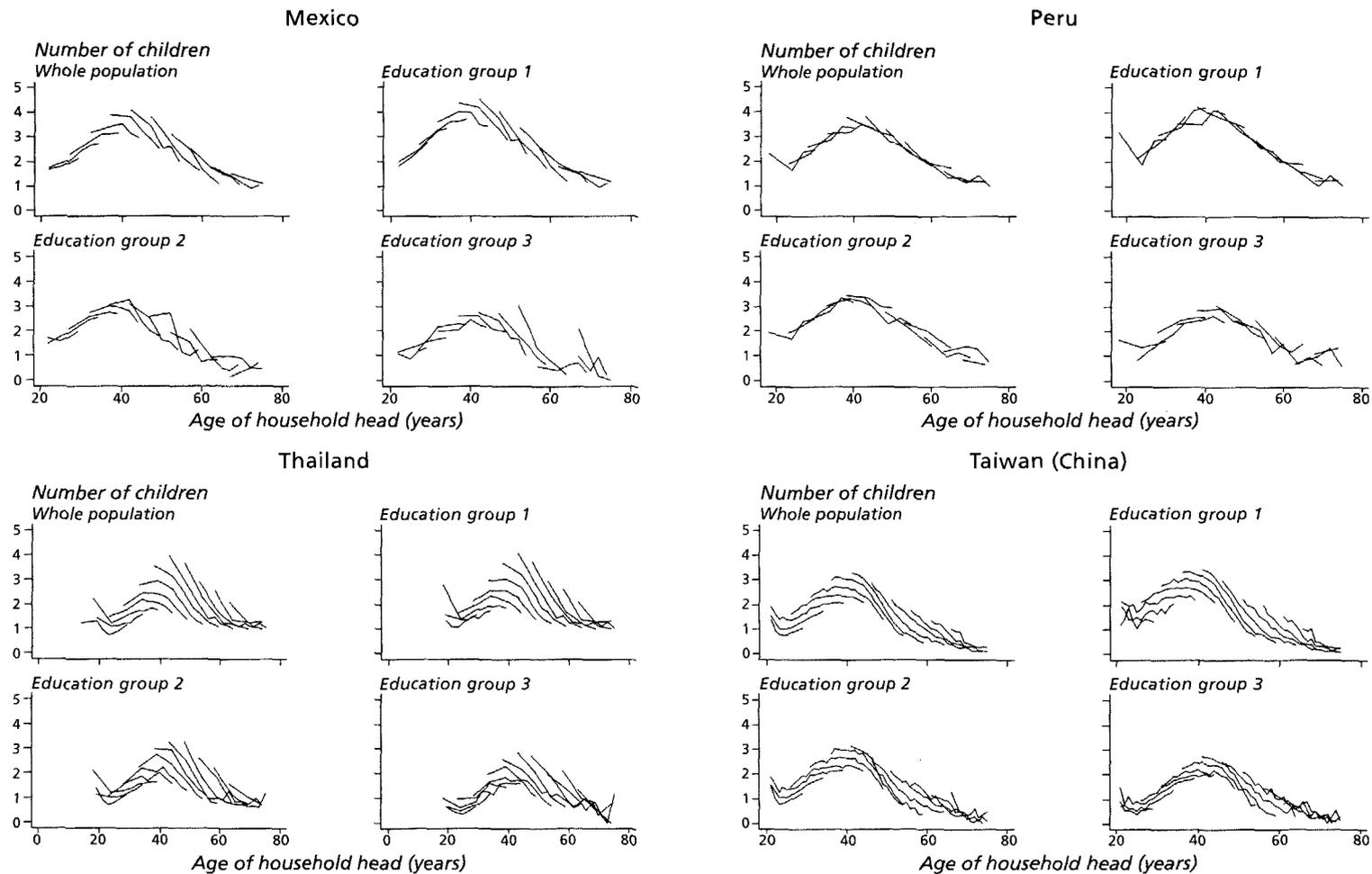
Source: Authors' calculations based on household survey data.

Figure 5. Size of Household, by Household Head's Age, Cohort, and Education Level



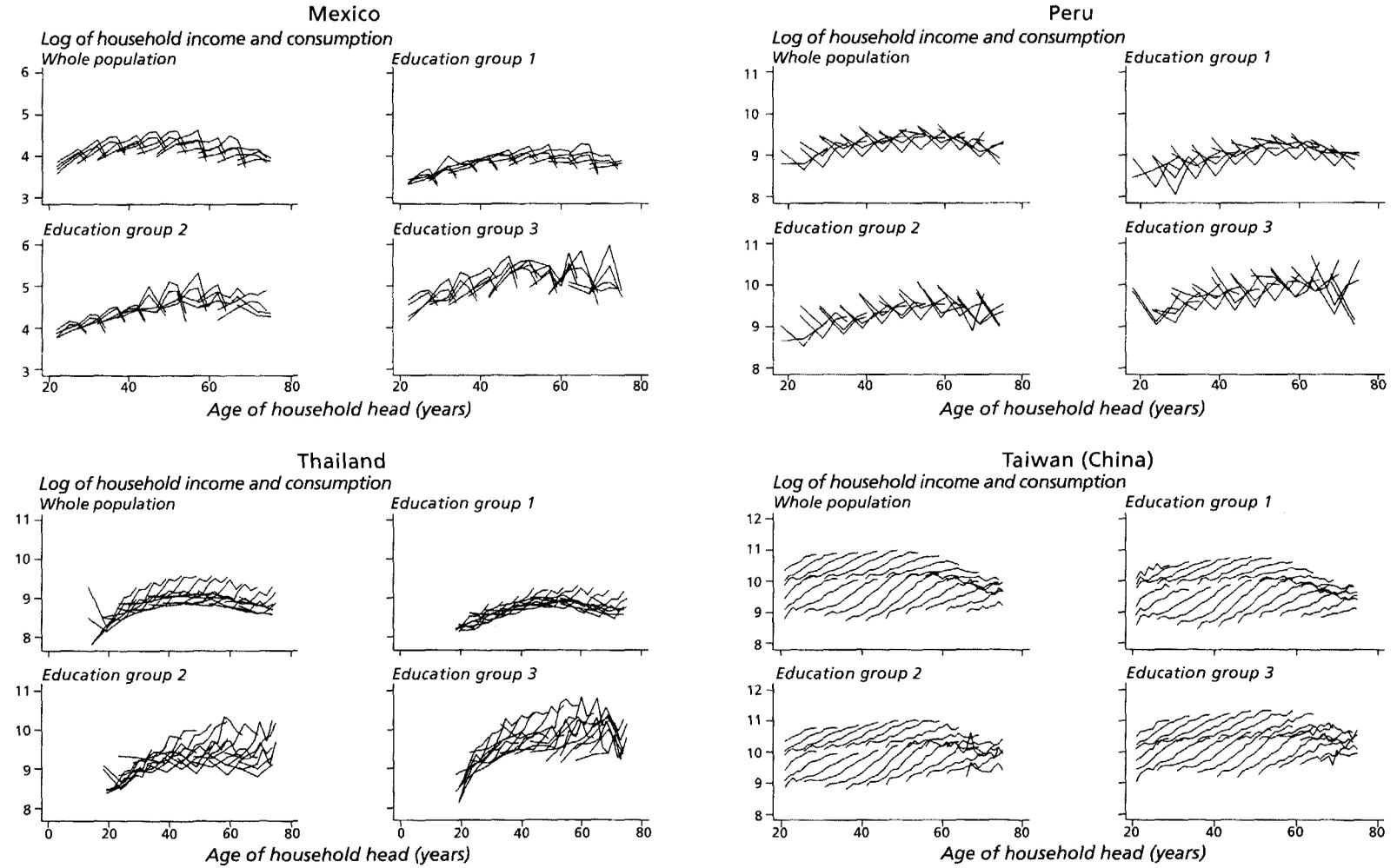
Source: Authors' calculations based on household survey data.

Figure 6. Average Number of Children, by Household Head's Age, Cohort, and Education Level



Source: Authors' calculations based on household survey data.

Figure 7. Disposable Household Income and Consumption, by Household Head's Age, Cohort, and Education Level



Source: Authors' calculations based on household survey data.

steeper over the life cycle. This difference is particularly apparent in Mexico and Thailand. Second, the experience of the two Asian economies, particularly Taiwan (China), is marked by the impressive growth of all cohorts. The entire life-cycle profile seems to shift up year on year starting in the early 1980s. Third, the experience of Peru, for which we have only four data points per group, is marked by the crisis of 1990, which is reflected in the 1991 survey. The decline in disposable income seems to have affected all education groups, and it seems more pronounced for the youngest cohorts. In comparison with this drop in income, even the declines observed in Mexico after the 1995 crisis (reflected in the 1996 survey) look small. In Mexico, unlike in Peru, the decline is more apparent for the most educated group.

In figure 8 we plot the log of per capita total and nondurable consumption. Overall, the per capita profiles are flatter than the corresponding household graphs in figure 7, although strong aggregate growth masks this fact in Taiwan (China). This result is consistent with the evidence reported by Attanasio and Browning (1995) for the United Kingdom and by Attanasio (1994) for the United States. An important difference between the regions emerges from figure 8: while the 1995 crisis in Mexico and the 1990 crisis in Peru are evident in the data, the data for East Asia are much smoother.

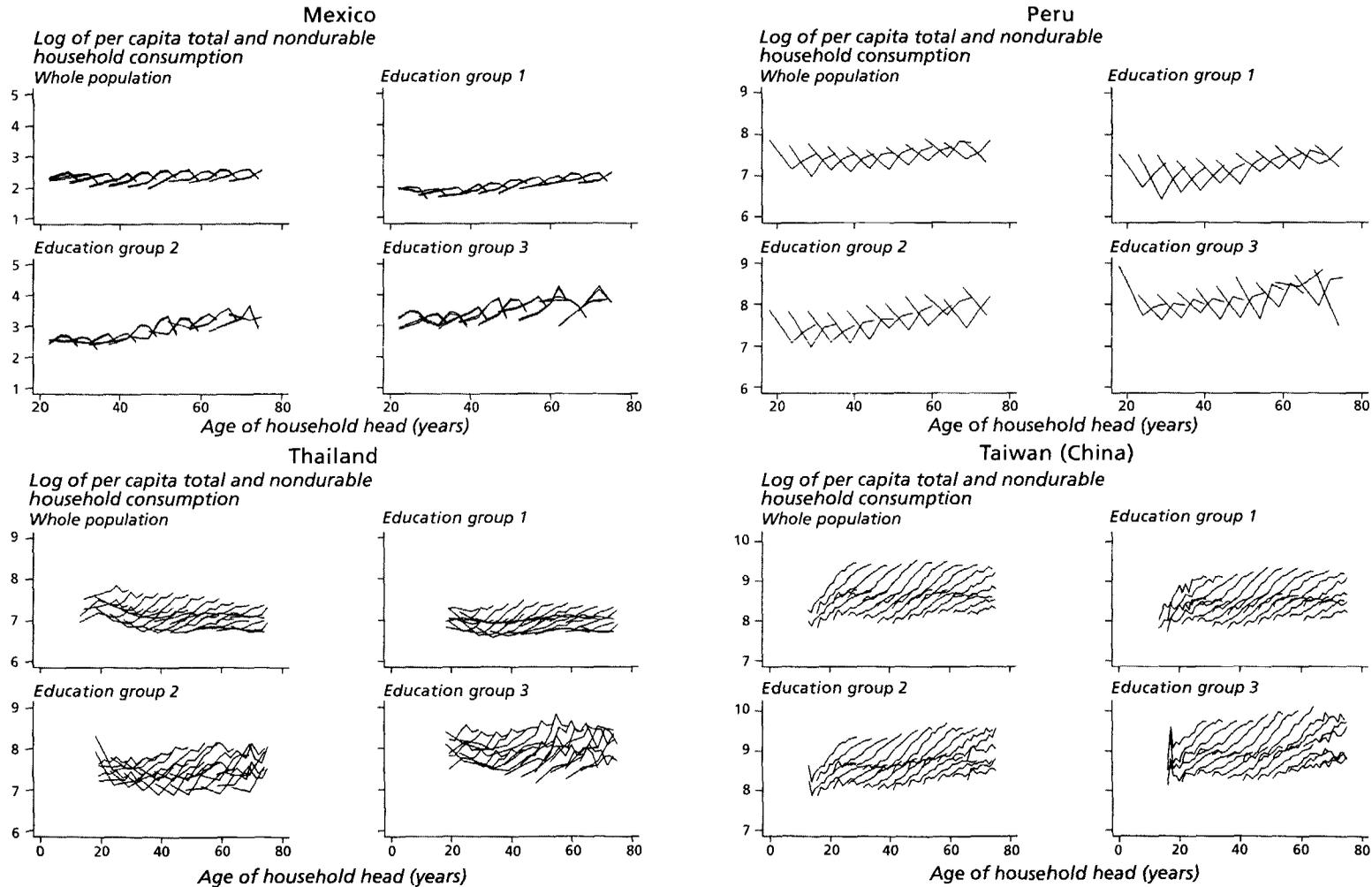
In figure 9 we plot two definitions of savings rates. The first (s_1) includes all expenditure items, while the second (s_2) excludes durable goods from the definition of consumption. The shape of the profiles is roughly similar, regardless of the definition used. In Mexico and Peru more educated households do most of the saving. This feature is not inconsistent with the life-cycle model, as the more educated face a steeper income profile. There seems to be no strong tendency for life-cycle profiles of saving to decline (and become negative) in the last part of the life cycle. The differences between the two definitions are greatest in Taiwan (China), where the hump shape is also more apparent.

As Deaton and Paxson (2000) recently emphasized, one reason that a clear hump shape consistent with the life-cycle hypothesis is not observed may be the changes in household composition documented earlier. Individuals may behave as the theory predicts, but aggregation into households that change in size, composition, and needs may mask these shifts in behavior. Deaton and Paxson suggest a method for identifying individual savings profiles from household data under some assumptions, and they conclude that a clearer hump is observed for individuals than for households in Taiwan (China) and Thailand.

Labor Supply and Wages

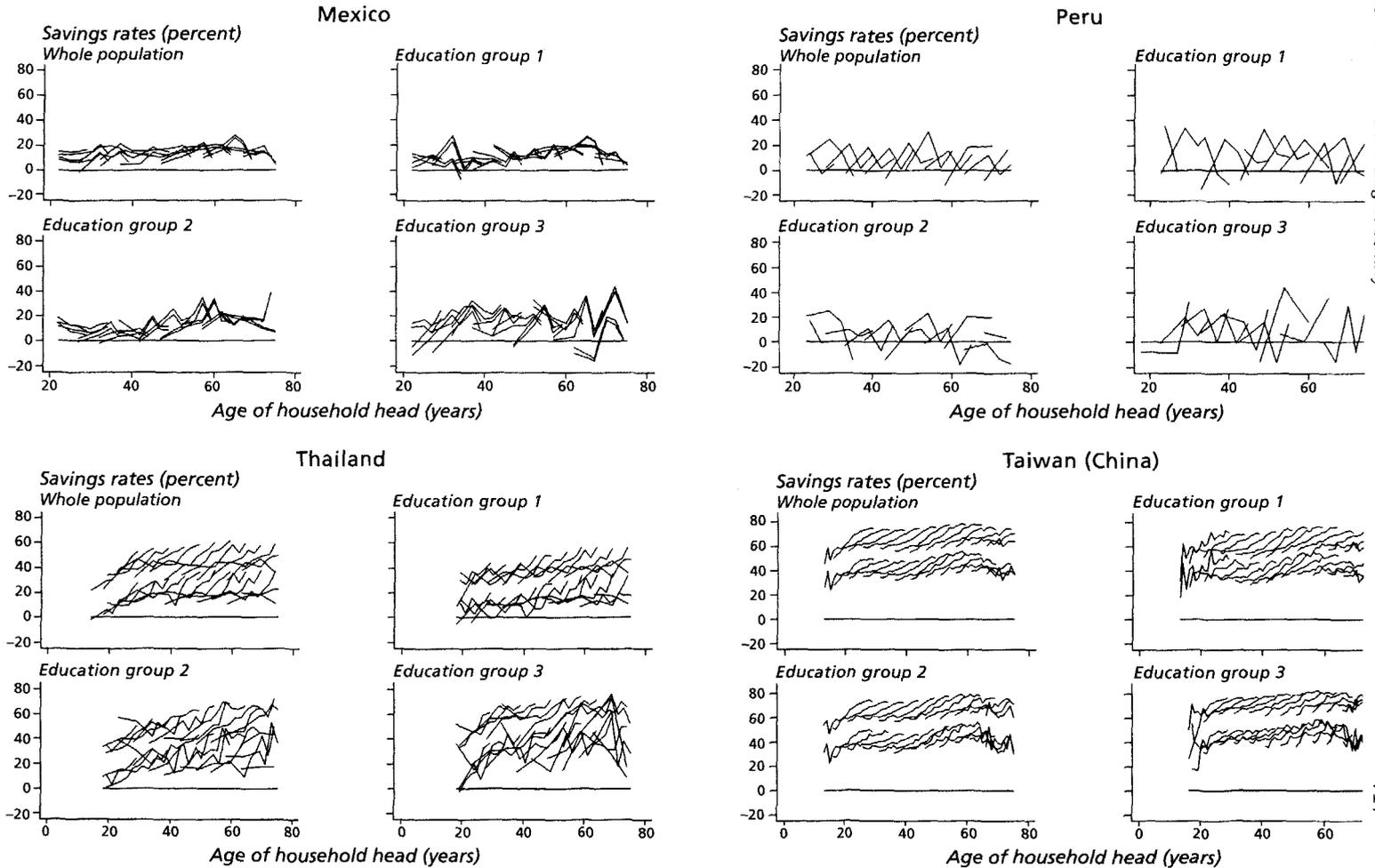
Households' capacity to save is determined largely by their income. Households' labor income reflects the wages paid in the market and labor force participation rates. In figure 10 we plot male wages for Mexico, Peru, and Taiwan (China) (Thailand is not presented because of data limitations). The results are unsurprising. First, the profile for more educated individuals is not only higher but also much steeper, especially in Mexico and Peru. Second, for all education groups the effects

Figure 8. Per Capita Total and Nondurable Household Consumption, by Household Head's Age, Cohort, and Education Level



Source: Authors' calculations based on household survey data.

Figure 9. Household Savings Rates, with and without Durables, by Household Head's Age and Education Level



of the 1995 crisis are apparent in Mexico, and a strong negative effect from the 1990 crisis is apparent in Peru in 1991. The patterns for Taiwan (China) are much smoother and show a continuous increase.

Perhaps most interesting is that these income profiles for individuals are much more hump shaped than the household profiles in figure 7. This is consistent with the argument by Deaton and Paxson (2000) that even though individuals may behave according to the life-cycle model, their behavior may be blurred by aggregation into households.

In figure 11 we plot male and female labor force participation rates by education level. The results indicate strong differences between men and women, especially in groups with less education. Relatively large cohort effects are visible in the labor force participation of women in the least educated households. For men, retirement is much more gradual among the least educated men, especially in Mexico, Peru, and Thailand, than among the more educated who tend to retire around the same age. This phenomenon may be related to the fact that more educated individuals are more likely to work in the formal sector and therefore to be covered by pensions. Higher labor force participation rates enhance savings capacity, so these results are compatible with the apparent cohort effects in saving presented earlier.

Female labor force participation has several interesting features. First, it is much higher in Thailand than anywhere else. Among educated Thai women, labor force participation is essentially equal to that of men. Second, female labor force participation seems high in Peru, at least by Latin American standards. Third, there are strong cohort effects in Mexico, especially for educated women. Finally, Taiwan (China) is the only economy in which labor force participation seems to decline in the most fertile years of women's life and then to recover. This dip seems to be absent elsewhere and is much less pronounced for educated Taiwanese women.

The evidence on female labor force participation should be interpreted with care, as cross-country differences might reflect differences in the way survey questions are asked. But it is important to consider the effect that female labor force participation might have on measured saving, as it is typically linked to a substitution of market goods for home production. Moreover, female labor supply could also have the effect of diversifying risk and therefore reducing the incentive for precautionary saving. But there are also reasons that households in which the wife works may save more. If female labor force participation is temporary, for example, it makes sense to save more to smooth income over time.

Another key determinant of saving is access to pension benefits. Individuals who believe that they will receive a pension after retirement have less incentive to save. Because household surveys contain only limited information on this issue, we are unable to document differences among the economies in detail. But broadly speaking, the evidence is consistent with some general differences between East Asia and Latin America. While most Latin American countries (including Mexico and Peru) have long traditions of providing pension benefits in the formal sector, such benefits have not been standard in Asia (Lora and Pagés 2000). In Thailand compulsory retirement benefits were introduced only in 1999. Thus in addition to the lower

Figure 10. Male Wages, by Age, Cohort, and Education Level

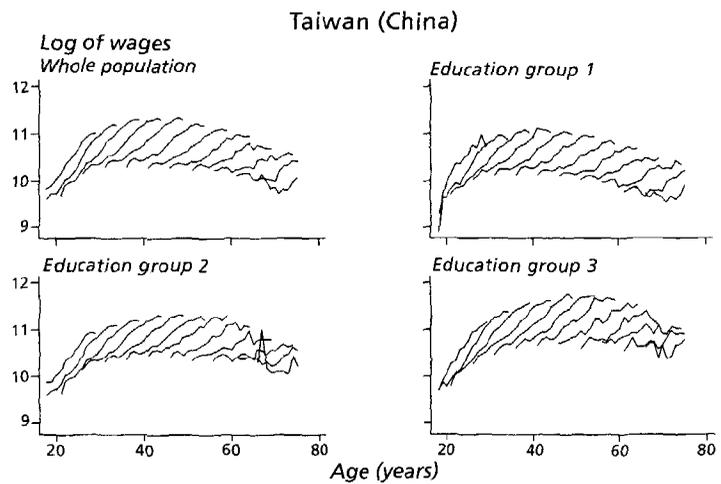
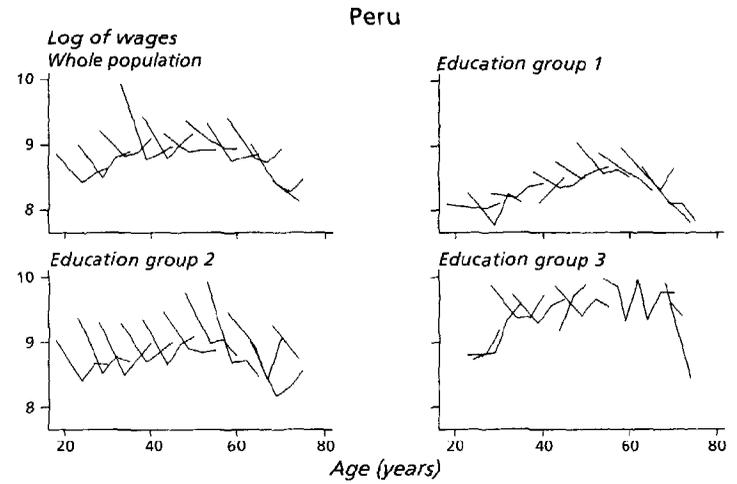
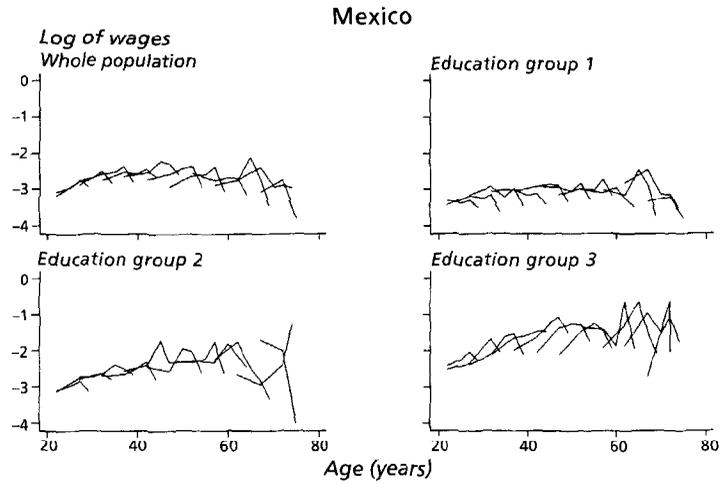
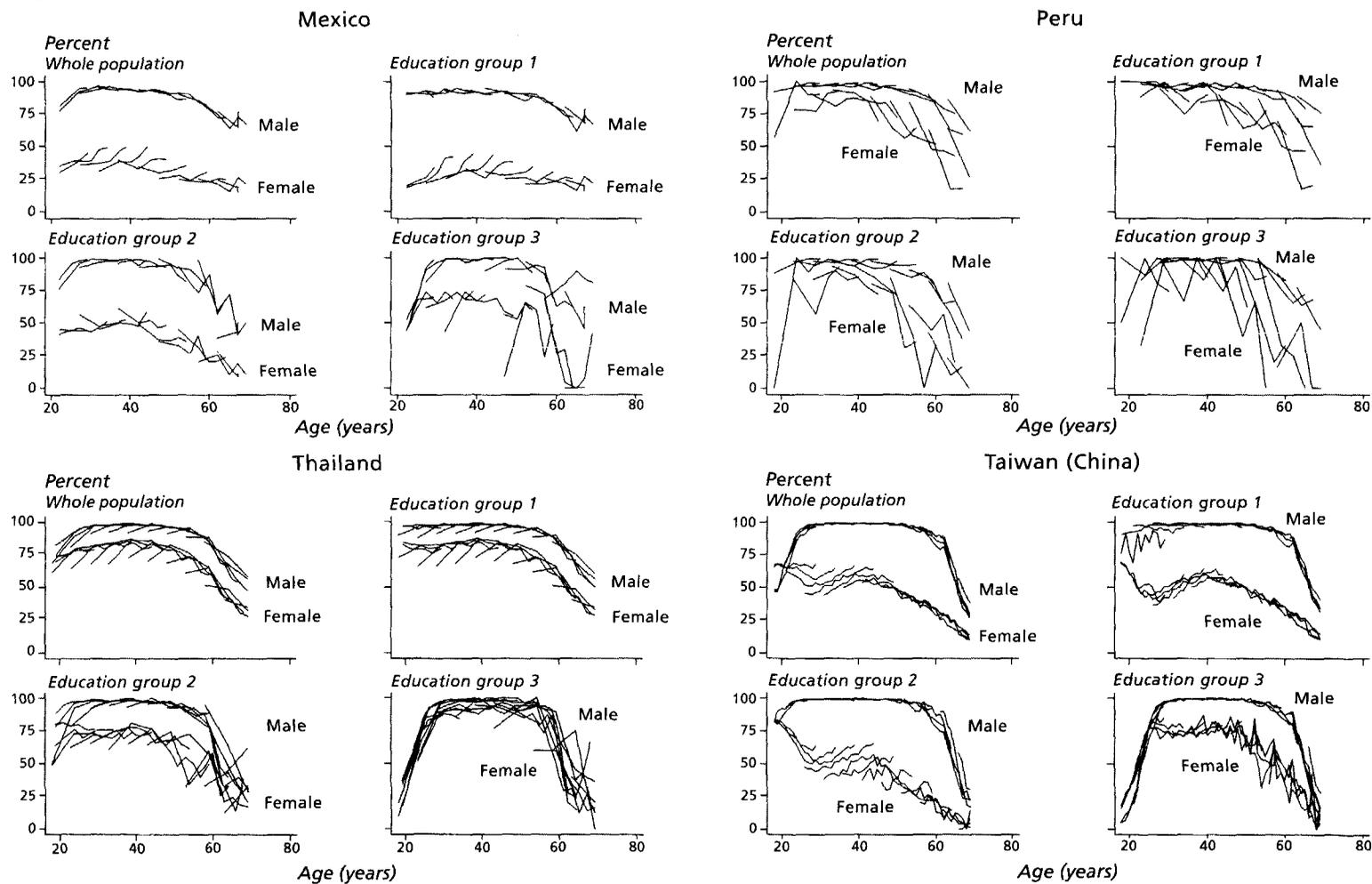


Figure 11. Male and Female Labor Force Participaton Rates, by Age, Cohort, and Education Level



Source: Authors' calculations based on household survey data.

wage levels and female labor force participation rates in Mexico and Peru, the fact that formal jobs in these economies have traditionally provided pensions may be another explanation for the lower household savings rates in the two Latin American economies than in the two East Asian ones.

Smoothing Savings Profiles

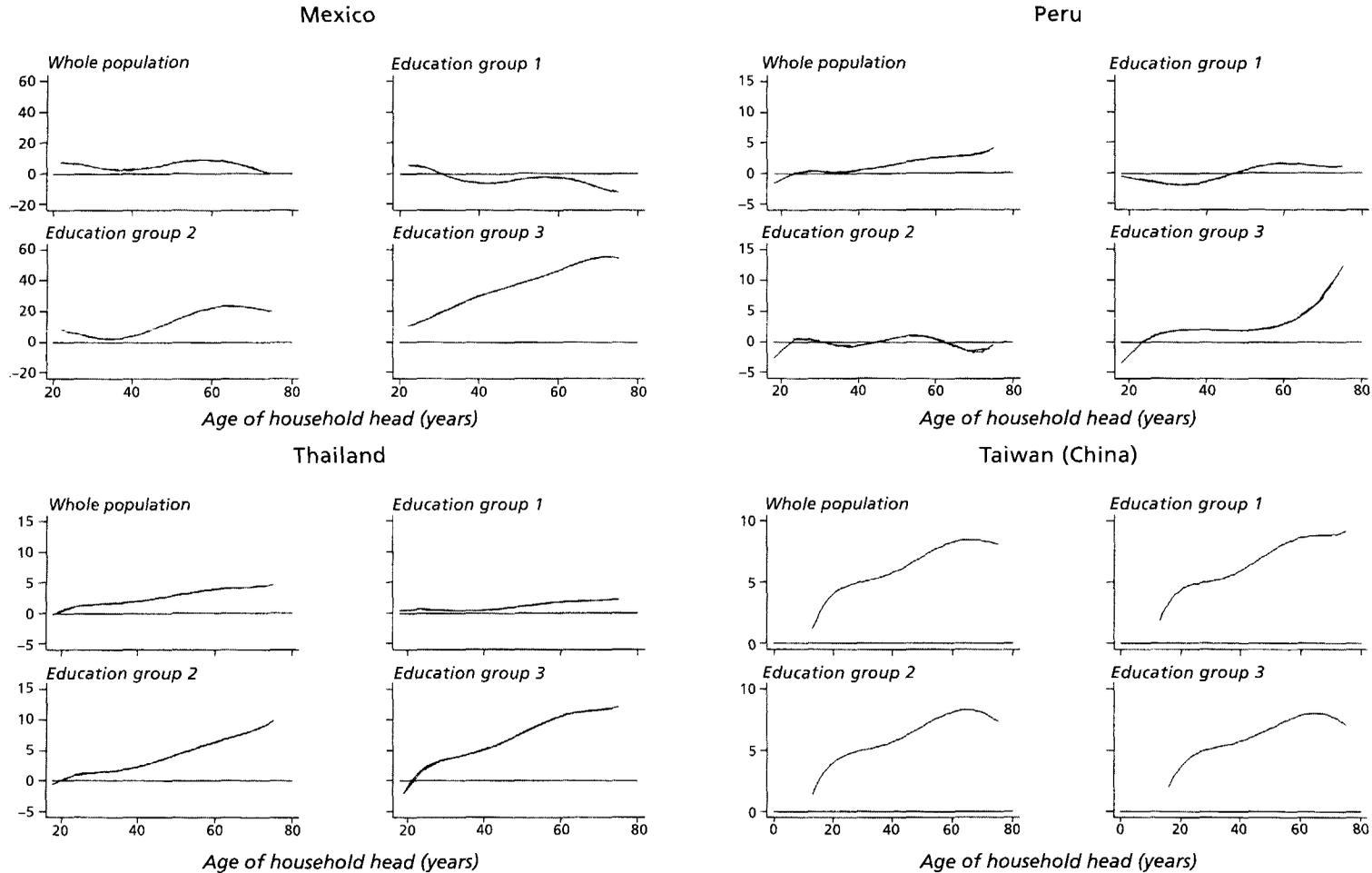
To identify the age profile of savings rates, we smooth the savings profiles presented in figure 9. However, as discussed earlier, one cannot separately identify time, age, and cohort effects without additional restrictions. We therefore regress the data points plotted in figure 9 on an age polynomial, a set of dummy variables for cohorts, and a set of dummy variables for years constrained to have zero mean and to be orthogonal to a time trend. The restriction we impose on the data is equivalent to assuming that all the deterministic trends in the savings rate data originate from a combination of cohort and age effects (see Heckman and Robb 1987, MaCurdy and Mroz 1995, Attanasio 1998, Deaton and Paxson 1994a, and Paxson 1996). We also assume that the age profile of saving is the same across cohorts, except for an intercept shift. While this assumption is not necessary for identification, it is forced on us by the fact that at least for Mexico and Peru, each cohort is observed for only a few years.

In figure 12 we plot the polynomial for an arbitrary cohort. Given our assumptions and restrictions, in none of the economies or education groups considered does the life-cycle profile have a marked hump shape. The profile for least educated individuals is very flat in Mexico, Peru, and Thailand, but not in Taiwan (China). For the most educated, the profile increases monotonically with age in Mexico, Taiwan (China), and Thailand and increases toward the end of the life cycle in Peru.⁶ Only in Taiwan (China) and in the middle group in Mexico and Peru do the smoothed profiles resemble the hump-shaped profile implied by some versions of the life-cycle model.⁷ Even for these groups, however, the decline starts only after age 65 and is very gentle—surprising in light of the reductions in labor force participation documented earlier. Regardless of whether the evidence is consistent with the life-cycle model, the lack of a hump in the middle of the life cycle has important implications for the effect of projected demographic trends on aggregate savings rates, as we show below.

In figure 13 we plot the dummy variables for the cohorts obtained from the same regression. The pattern indicates substantial differences across cohorts. In particular, the youngest cohorts (those with the lowest cohort numbers) seem to have much higher savings rates than older cohorts, although in Mexico this effect is not observed for the population at large.

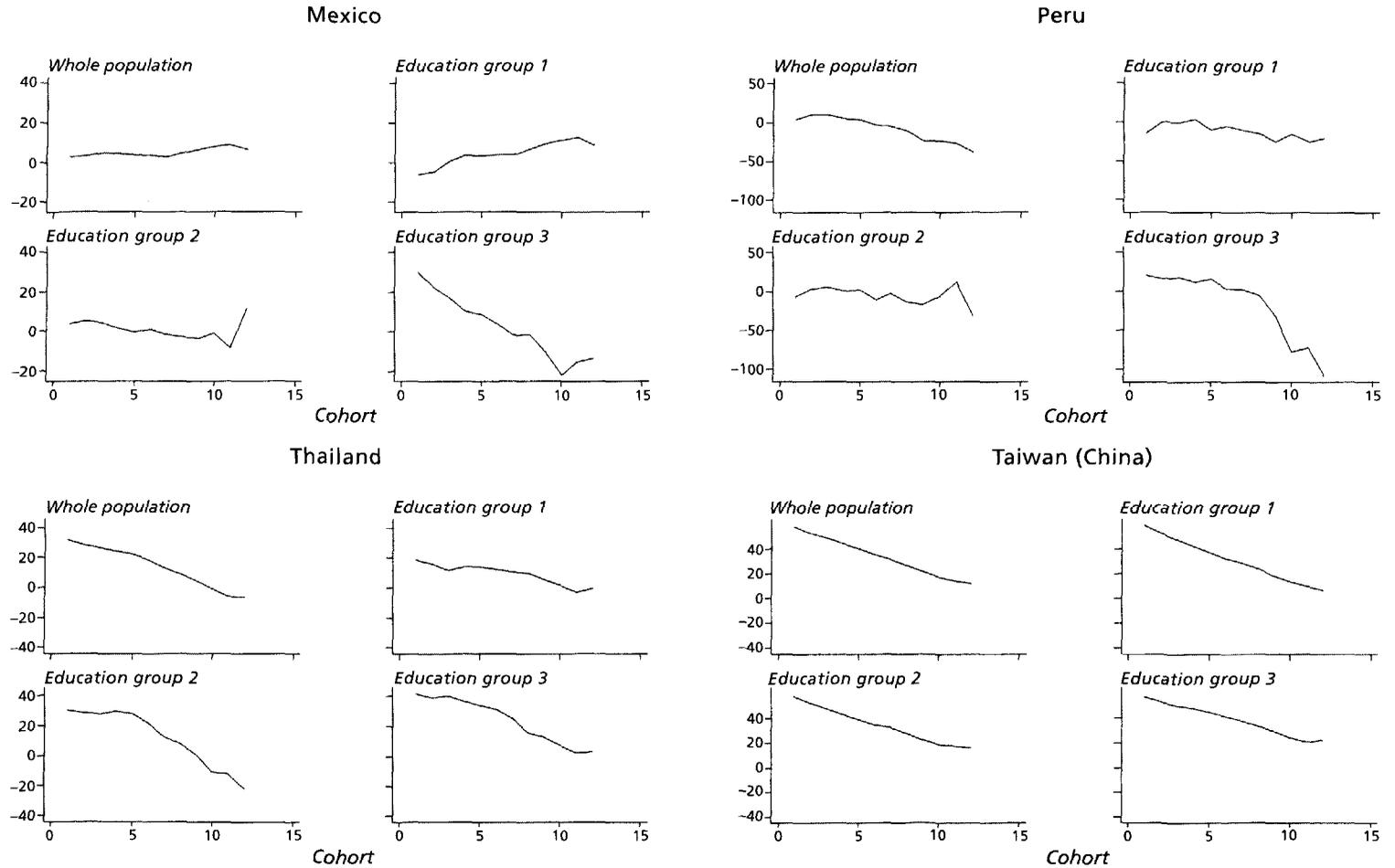
There are also important differences across education groups. In Mexico and Peru positive cohort effects are observed only for the middle group and the most educated. In Thailand there are positive cohort effects in all groups, but they are stronger for the middle group and the most educated.⁸ Taiwan (China) is the only economy in which differences across groups are not apparent.⁹

Figure 12. Age Effects on Household Saving Rates



Source: Authors' calculations based on household survey data.

Figure 13. Cohort Effects on Household Saving Rates



Note: Cohorts at the low end of the range are younger; those at the high end older.
Source: Authors' calculations based on household survey data.

Figure 14 plots the restricted year effects estimated together with the age profile plotted in figures 12 and 13. A large negative shock is evident for all education groups in Mexico in 1996 and in Peru in 1994. In Mexico the effect of the 1996 shock on the most educated is very large and negative, while the effects in previous years are all mildly positive (or very small and negative). For the least educated the aggregate shock is negative in 1992, and the positive time effects in the years before 1996 are smaller than for the rest of the population. The 1996 effect is similar for all groups.

The results for Peru are surprising in that in 1991, the year after the main crisis, the estimated residual is positive. This is a direct consequence of the fact that consumption falls more than income in the survey data. All groups experience a strong negative shock in 1994; only the most educated have a positive time effect in 1997. In Thailand time effects are smaller and stronger for the least educated. In Taiwan (China) the effects are very small, and time effects vary little across education groups.

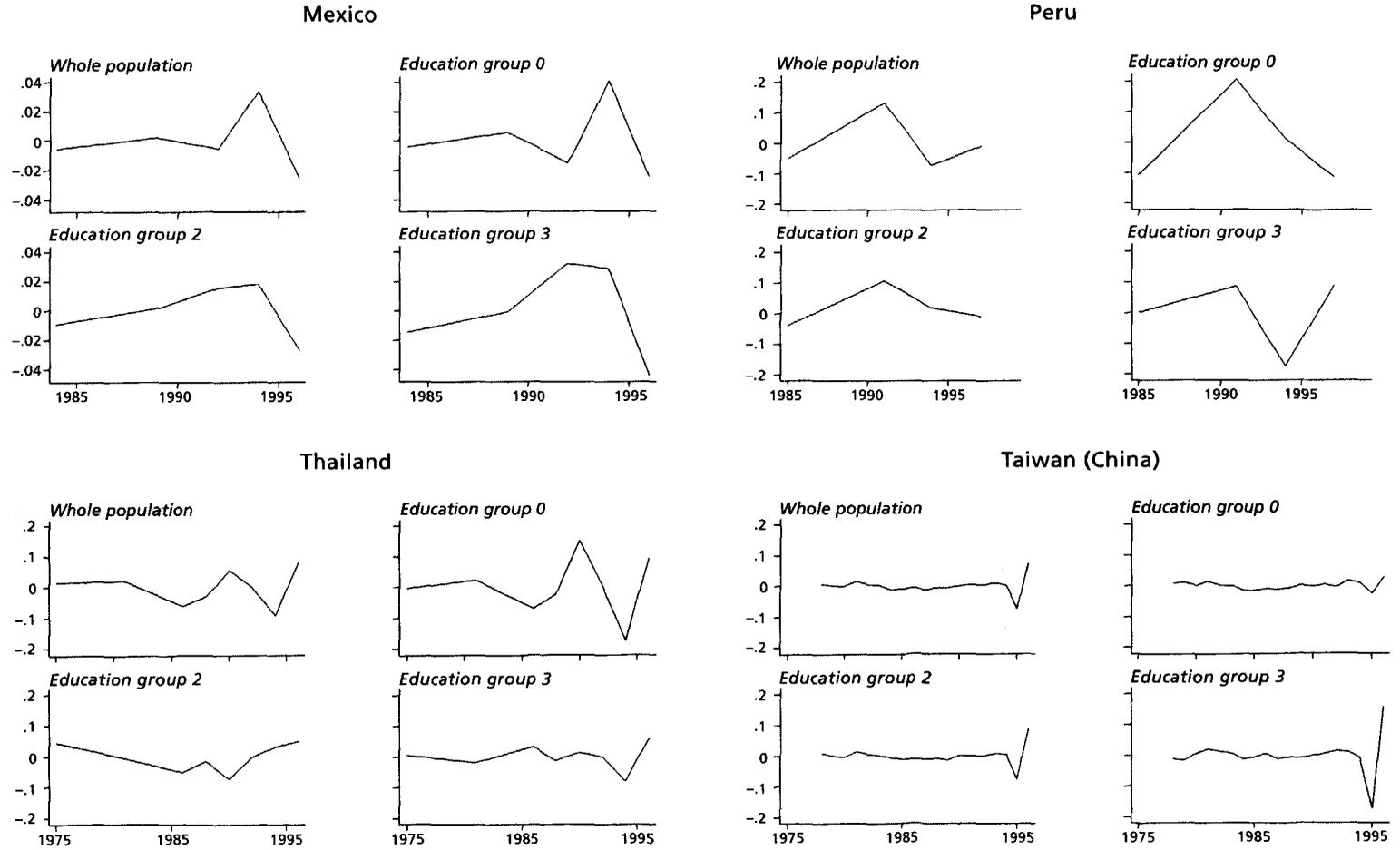
While Mexico and Peru experienced severe economic crisis during the period under analysis, the economic environment in Taiwan (China) and Thailand was much more stable. Households in Latin America have been exposed to an economic environment characterized by the use rather than the accumulation of savings, which seems to be one reason that they have lower savings rates and that savings have increased much less than in East Asia.

The age and cohort profiles estimated so far depend on the arbitrary normalization that year effects sum to zero and are orthogonal to a linear trend. An alternative restriction that allows the identification of age profiles is the assumption that there are no cohort effects in savings rates. This would be the case if cohort effects in income and consumption exactly canceled each other out. Deaton and Paxson (1994a) suggest that some versions of the life-cycle model do imply such a restriction. By imposing it, we can identify unrestricted year effects.

We plot these age profiles in figure 15, using the first year of the sample as the intercept. For the entire sample the assumption of no cohort effects increases the size of the hump in all economies, with savings rates peaking just after age 60. This pattern is roughly consistent with the life-cycle model. For the least educated the effect is roughly similar: a profile that looked basically flat now reveals a modest hump, with a peak just past age 60. The profile for the middle education group is almost unaffected. The largest effect is observed for the most educated group. The assumption of no cohort effects implies a hump-shaped profile with a dramatic decrease after age 65 in Mexico, Taiwan (China), and Thailand, while the previous profile increased monotonically with age. In Peru there is a clear hump shape, with an increasing trend after age 60.

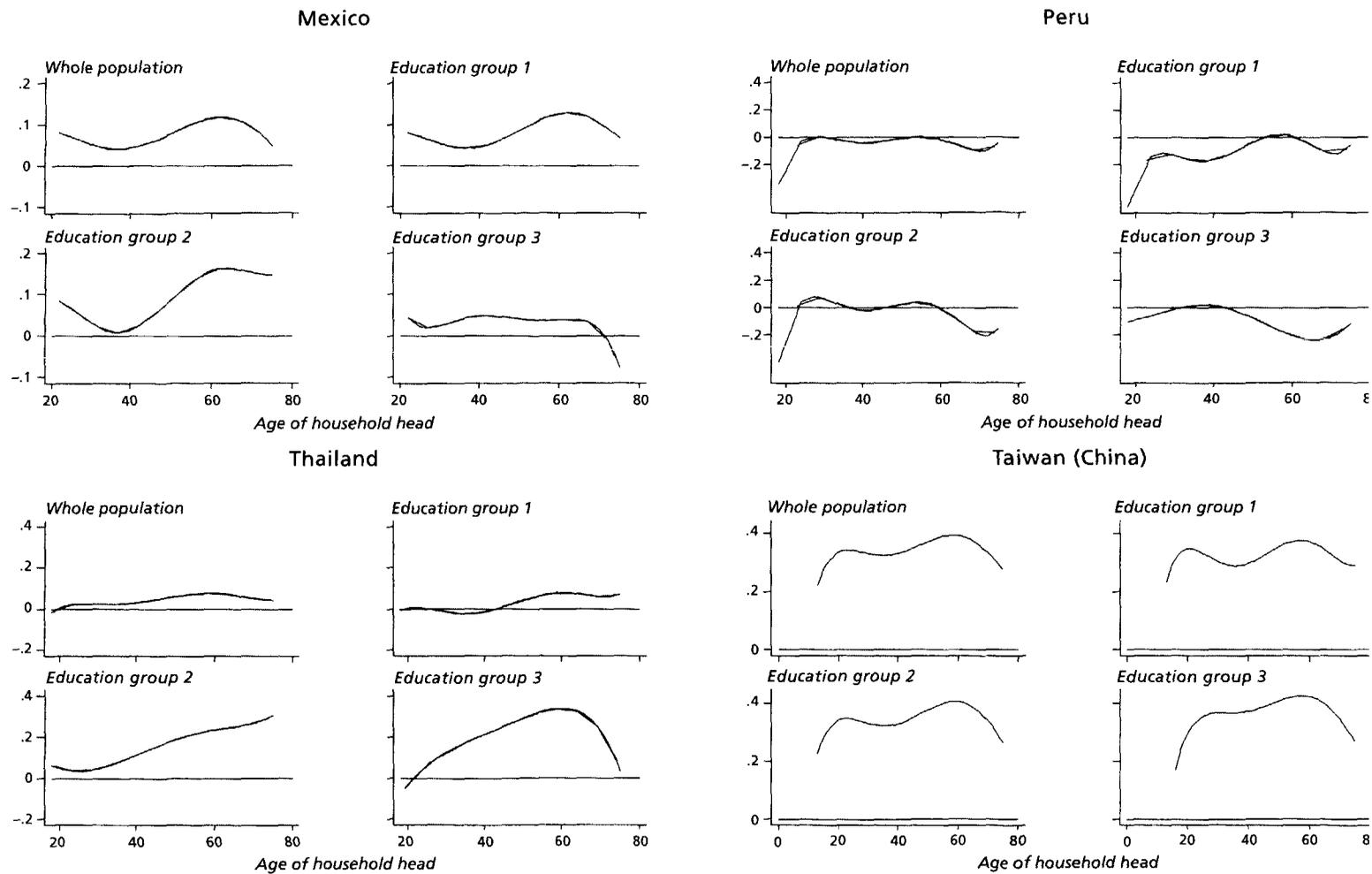
The unrestricted time effects in figure 16 are also different. These dummy variables for time now capture all the trends in the data that were previously interpreted as cohort effects. In Mexico the strongest negative effect in 1996 is observed among the least educated household heads. Time effects are also nega-

Figure 14. Year Effects on Household Saving Rates



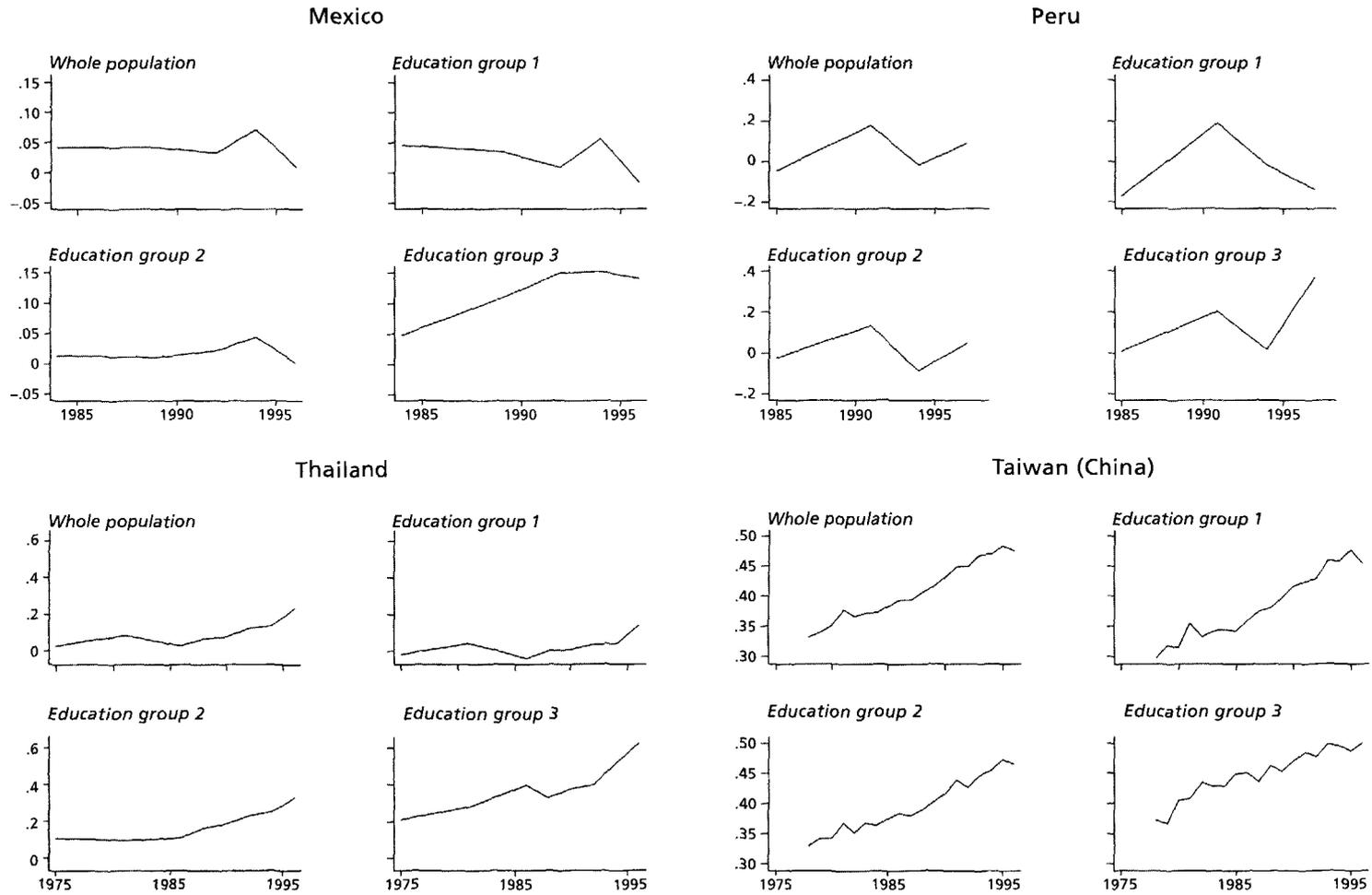
Source: Authors' calculations based on household survey data.

Figure 15. Age Profile of Household Saving Rates, No Cohort Effect Assumption



Source: Authors' calculations based on household survey data.

Figure 16. Year Effects on Household Saving Rates, No Cohort Effect Assumption



Source: Authors' calculations based on household survey data.

tive in 1996 for the intermediate group; surprisingly, the most educated show a positive time effect in 1996. This result may indicate their greater capacity to smooth variations in income. Similar results are obtained for Peru in 1997. In Taiwan (China) and Thailand the strong cohort effects are mirrored in strongly increasing time effects.

Savings Profiles in East Asia and Latin America

Given the exercise we propose in the next section, it is worth comparing the demographic factors and smoothed savings profiles of East Asia and Latin America. The demographic factors confirm the results of the aggregate data: the demographic transition is much more advanced in East Asia, where fertility rates are much lower and families smaller. The size of the cohort effects in household size and number of children, however, suggests that Mexico and Peru will be in a similar position in a few years.

The picture emerging from the analysis of savings behavior is more complex and more difficult to interpret, if for no other reason than that the estimated effects are conditional on strong and untestable identification assumptions. If we assume that all the trends in the data are explained by age and cohort effects and compare the age profiles of aggregate saving, in Mexico we find a mild hump shape with a peak around age 60, while in Peru and Thailand the profiles rise steadily with age. This picture hides strong differences among education groups, however. In these three countries more educated individuals seem to show no tendency to reduce savings in the last part of the life cycle. In Thailand similar patterns emerge for the other two groups, although the profile for the least educated is basically flat rather than rising. In Mexico and Peru the two lower groups show a mild hump in the last part of the life cycle. This surprising result presumably occurs because although the most educated do most of the saving, their population weight is much lower.

If we move on to cohort effects, the differences between the two regions are even more apparent. In Taiwan (China) and Thailand the two groups with more education show strong positive cohort effects; in Mexico and Peru a similar pattern is observed only for the most educated. At the other extreme, the least educated have negative cohort effects in Mexico and no cohort effects in Peru, while in Taiwan (China) and Thailand they have mildly positive cohort effects. Time effects, as expected, also differ markedly. Strong negative shocks were experienced in Mexico in 1996 and in Peru in 1994; in Taiwan (China) and Thailand time effects were mostly positive. The effects are reinforced when cohort effects are assumed to be equal to zero. Interestingly, time effects in both Latin American countries are stronger (more positive) for the most educated.

The picture changes, however, once we use a different identification assumption—the absence of cohort effects. Under this assumption the estimated profiles do exhibit a hump in the middle of the life cycle, with a peak just before retirement. While this identification assumption is questionable, the results, which

indicate a slight hump in the middle of the life cycle, are not inconsistent with the life-cycle model.

Projections

Is the gap in aggregate saving between East Asia and Latin America likely to narrow in the future, or will it continue to expand? For believers in a standard version of the life-cycle model, there are reasons to believe that the gap will narrow. Latin America is on the verge of a rapid demographic transition that will result in a shift in the population toward age groups expected to save more. Although less pronounced, the predicted demographic changes are similar to those that preceded the boost in private savings rates in East Asia.¹⁰ Whether this expectation will be fulfilled depends largely on the shape of the age profile of saving and how it evolves in the future.

The basic idea we investigate in this section is as follows. If one believes that life-cycle saving is important and that saving is concentrated in a certain age group, then an economy or region in which the share of the population in that age group increases because of the demographic transition (neglecting possible general equilibrium effects on factor prices) might experience an increase in aggregate saving during that transition. Because it is plausible that most life-cycle saving occurs in the middle of the life cycle, economies that experience an increase in the share of middle-aged individuals (and a decline in both youth and old age dependency ratios) might also be expected to experience a temporary increase in saving. This argument is relevant for Latin America, since that region is about to enter such a demographic phase. We attempt to quantify these effects, using the evidence on savings behavior discussed in the previous section.

The projections we present should be interpreted with care. The simulations are based on simple reduced-form relationships, and their use in forecasting future aggregate savings rates assumes that they are stable over time. Specifically, we need to assume that the savings profiles identified in the previous section (and the income profiles used to weight them) do not change in the future. Changes in age profiles of saving could be induced by changes in the shape of the life-cycle profiles of earnings, by changes in factor prices (wages and interest rates), and, in models with habit formation, by the process of growth. Moreover, in some cases we use the last year of our sample as a benchmark, even though it may not be representative.¹¹ These caveats notwithstanding, the exercises proposed should serve as benchmark calculations for quantifying the potential effects of demographic trends on aggregate saving given the evidence on life-cycle saving.

One important difference between our analysis and earlier studies that have used either macroeconomic aggregate data (such as Behrman, Duryea, and Székely 1999b) or simulations of general equilibrium models (see Attanasio and Violante 2000) is that while we focus only on household saving, those studies focus on domestic saving, which includes public and firm saving. Moreover, those studies are more likely to take into account pension assets and liabilities. In some sense, then,

the exercise based on microeconomic data is more limited in scope. As noted, however, even if the analysis with microeconomic data excludes some important elements of household saving, it also brings considerable gain in that it allows identification of some of the mechanisms driving the dynamics of saving.

Forecasting Household Saving with Microeconomic Data

To forecast aggregate household savings rates, we use the accounting identity given by equation 1, the smoothed profiles we estimated earlier, and demographic projections. Specifically, for any year beginning in the late 1990s, we compute:

$$(1) \quad \hat{s}_t^{ag} = \sum_c \hat{s}_t^c \frac{\hat{Y}_t^c N_t^c}{\sum_c \hat{Y}_t^c N_t^c}$$

where \hat{s}_t^c is the savings rate of group c at time t predicted by the smoothing procedure used to produce figure 12. In particular, we use the estimated age profile (if c is the year of birth of a cohort, its age will be $t - c$ at time t) and the relevant cohort effects. We use an analogous procedure to compute \hat{Y}_t^c . That is, we estimate age and cohort effects using the same procedure used to identify the age and cohort effects for savings rates reported in figures 12 and 13. The N_t^c are from United Nations (1998).

We define the aggregate savings rate as the savings rate of households whose head is 23–75 years old. Because we forecast aggregate savings rates far into the future, new cohorts will join the sample and some cohorts will leave it. For the new cohorts we use the same age profile as for the other cohorts and the cohort effect of the youngest cohort in the sample. This exercise can then be extended to consider different education groups, by simply repeating the exercise for each education group and then aggregating across education groups given some projections about each group's relative size. Because we do not have forecasts of future generations' educational attainment, for future cohorts we use the proportions observed in the youngest cohort. This procedure ignores the fact that future generations are likely to be more educated. But ignoring education groups is equivalent to assuming that changes in the composition of future households will leave the shape of the life-cycle profile of savings rates unchanged, which is obviously unrealistic.

Our aim is not to reproduce the level of aggregate savings rates or efficiently forecast their evolution. As we discussed earlier, there are many reasons why microeconomic data do not match up exactly with aggregate statistics. These reasons are compounded by the fact that the shape of the life-cycle profile is likely to change as a result of changes in its determinants. Our more modest aim is to understand the implications of our estimated age profiles and the predicted demographic trends for the evolution of aggregate savings rates. The results of even this limited exercise must be interpreted with caution, however, for several reasons. First, as emphasized, we identify age and cohort effects under the arbitrary normalization that year effects have zero mean and no trend. Second, we assume that the age profile of savings rates

and income is the same across cohorts, except for an intercept shift—an assumption forced on us by the lack of a long time series of cross-sections. Third, even if existing cohorts (within an education group) have the same age profile of savings rates and income, it is likely that changes in wage profiles, household size and composition, labor force participation, and institutional factors will have an effect on the age profiles of saving. Fourth, changes in the stock of human and physical capital are likely to change wage and interest rates, inducing further changes in age profiles.

Evidence Based on the Microeconomic Simulations

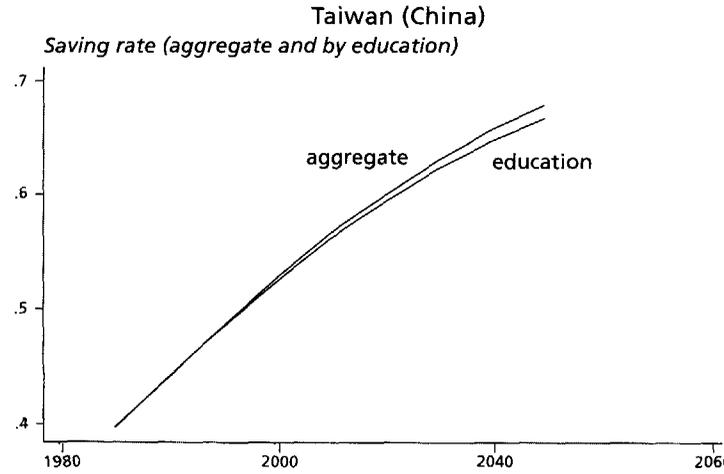
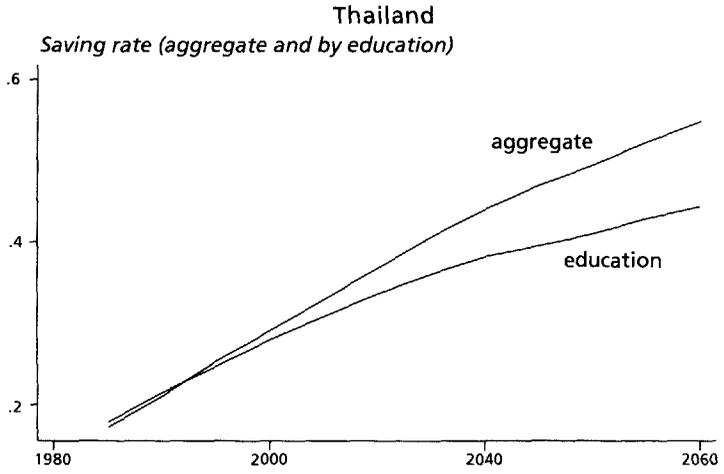
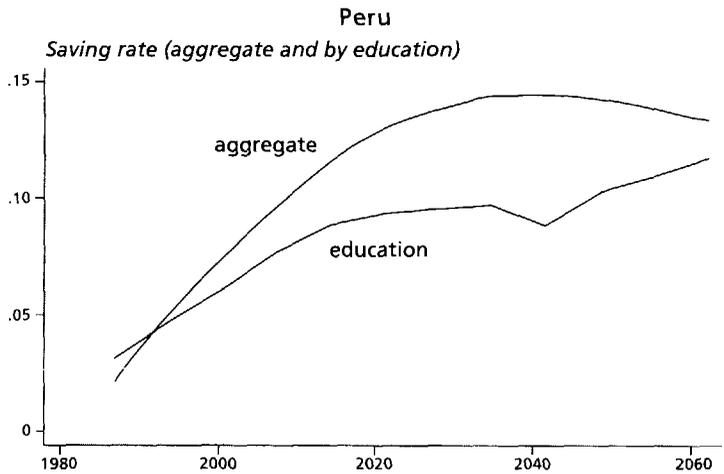
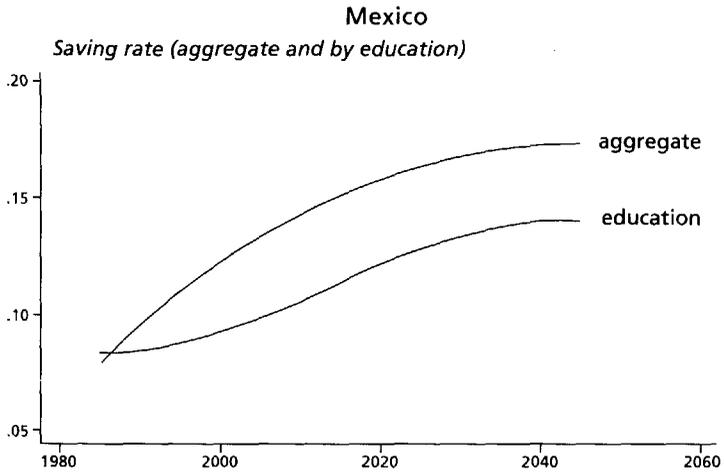
We start with the forecast for Mexico, performing two different exercises. In the first we use the age profile and cohort effects for the overall population (the top-left panel in figures 12 and 13). In the second we use education-specific profiles and cohort effects (the other three panels in figures 12 and 13). To aggregate across education groups, we use the proportions in the sample for cohorts currently alive and the proportions in the last cohort for future cohorts. We plot both sets of forecasts in the first panel of figure 17. Both forecasts show a marked increase in aggregate savings rates that starts leveling off only around 2040. The forecast that does not take into account the education split shows a larger increase, although the forecast that uses the education-specific profiles starts to catch up.

The exercise for Peru shows an increase similar to that for Mexico, although the forecast using education-specific profiles and cohort effects catches up more rapidly. The most interesting result, however, is that the forecasts for the two East Asian economies reveal even greater increases in saving than those for Latin America.

While at first sight these forecasts seem to support the hypothesis that demographic trends will lead to an increase in aggregate savings rates in Latin America, a more careful consideration of the mechanics behind the forecasts shows that this is not necessarily the case. Demographic shifts actually play a small role. The main reason for this is the lack of a hump in the age profiles of saving. Therefore, even when the share of the population between ages 40 and 60 increases—as it is projected to do in the next 40 years in Latin America—the change has little effect on aggregate savings rates; most of the increase in aggregate savings rates is driven by the cohort effect. In figure 12 the cohort effect is strongest for the youngest cohort. Because we give the intercept for that cohort to future cohorts that enter our computation, as older cohorts (with lower intercepts) disappear, the aggregate savings rate increases. The increase is smaller for the education-specific profiles because the cohort effect for the youngest cohort averaged across education groups is smaller than that estimated for the entire population. The catching up occurs because, as the population becomes more and more similar to the first cohort, it has both the intercept and the proportions of educational attainment of that cohort. Because the youngest cohort will be more educated, the savings rate increases slightly more quickly after the first few years.

Similar considerations hold for Taiwan (China) and Thailand, with two notable differences. First, the increase is more marked than in Mexico. And second, the forecast using education-specific profiles catches up faster. We should not read much

Figure 17. Simulated Saving Rates



Source: Authors' calculations based on household survey data.

into the first effect because the increase is driven mainly by the estimated intercept for the youngest cohort, as it was for Mexico. Moreover, the absolute level of these profiles does not have a straightforward interpretation because of the definitional and measurement issues raised earlier.

Because the projected increases in aggregate savings rates are driven by the estimate of a single parameter, these results should be interpreted with extreme care. Nevertheless, given our identification assumptions, younger cohorts do seem to be saving more today than their predecessors. Whether this pattern can be maintained in the future will depend on the evolution of the determinants of savings.

We showed earlier that the shape of the estimated profiles depends strongly on the assumption made to identify the age profile of savings rates. If we assume that all the trends observed in the data come from either cohort or age effects and that year effects have zero mean and are orthogonal to such trends, there is no strong evidence of a hump-shaped savings profile in Mexico, Peru, or Thailand. But if we assume that there are no cohort effects, so that year effects can be estimated in an unrestricted fashion, the Mexican and Peruvian age profiles show a marked hump, as do those for the most educated households in Thailand. Both sets of profiles can be used to forecast future household savings rates to check the extent to which the demographic transition is likely to affect aggregate household savings rates.

We perform the exercise described above for Mexico, using the numbers plotted in figure 15 as the age profile of savings rates. For the population as a whole and for two of the three education groups, the profiles show a marked hump. This approach, however, assumes that there are no cohort effects, shutting down the main source of increase in the aggregate savings rate in figure 17.¹² The result we obtain is that savings rates start to decline around 1995, continue to decline for about 20 years, and then begin to increase around 2020 as the population share corresponding to the hump in the savings rate profile starts to increase. But the increase in savings rates is minuscule (about 0.2 percentage point). There are two reasons for the small size of this effect. First, the hump in figure 15 is not very pronounced. Second, although the demographic change we project is relatively large, it results only in a change in weighting, which cannot have a very large effect given the size of the hump. Although the scope of this exercise is limited, we can conclude that, given the estimated shapes of the age profiles of saving (and the projected demographic trends), these forces by themselves are unlikely to result in a large increase in aggregate savings rates if the current economic environment prevails.

Conclusion

In this article we compare East Asia and Latin America in an important dimension in which they have diverged markedly during the past three decades—household savings behavior. In addition to the life-cycle profile of household saving and other variables of interest, we characterize differences in savings behavior across education

groups in two economies in each region. The evidence indicates large differences, both across economies and, within each economy, across education groups.

We document the huge disparity in the level and growth of household saving between Mexico and Peru on the one hand and Taiwan (China) and Thailand on the other. Normalizing time effects to have zero mean and no trend, we identify cohort effects in the data and confirm that relative to older generations, younger generations in East Asia are saving much more than their counterparts in Latin America. Our analysis suggests three main reasons why cohort effects are stronger and total household savings are much higher in East Asia.

First, East Asian households, especially those in younger generations, have had greater capacity to save because of higher income growth, lower fertility rates, different household structure (with more elderly individuals living in extended households, which prevents us from observing a decline in saving toward the end of the life cycle of individuals), and a much more advanced demographic transition (with larger shares of the population in the ages at which productivity and saving peak).

Second, the macroeconomic environment in Latin America was highly volatile, and the two countries we analyze were subject to severe shocks during the period under study. In contrast, East Asia was much more stable economically. Thus in Latin America the context was one in which household savings were typically used rather than created, while in East Asia savings were built up smoothly to accumulate resources.

Third, in Latin America almost all household savings are generated by the richest 20 percent of the population, while in East Asia saving is much more widespread. As a result, in Latin America the savings rate of households in the richest quintile does not differ significantly from that of their East Asian counterparts, but there are huge disparities between the richest households and the rest of the population. For the poorest 50 percent of households in Latin America, savings have been much less responsive to increases in income during periods of economic growth and more sensitive to declines during downturns. Thus differences in the capacity to save across the income distribution account for an important part of the difference in total household saving between the two regions.

We do not find strong evidence of negative saving or even declining saving in the last part of the life cycle in any of the economies studied. While this finding contradicts a simple version of the life-cycle model, a conclusive judgment can be obtained only if one explicitly takes into account the variation in needs induced by changes in household size and composition over the life cycle as well as changes in labor supply behavior. In this study we document differences in life-cycle profiles in these variables, but we do not explicitly consider their effect on savings rates.

Another issue we have ignored is the effect of different institutional settings, particularly pension arrangements, on savings behavior over the life cycle. Using evidence on household saving to test alternative models of consumption and savings decisions does not obviate the need to consider these factors. However, some of the evidence—particularly that from Taiwan (China) and Thailand—suggests a need for a more complex model than the simple version of the life-cycle model. Paxson

(1996) and Deaton and Paxon (1994a) suggest considering models with habit formation, in which growth in itself induces an increase in savings rates, at least until the stock of habits “catches up.” This issue is extremely important for understanding the relationship between growth and saving, and more detailed study is needed.

In all four economies we find that households headed by the most educated save considerably more than households headed by those with less education (the exception could be Taiwan [China]). This finding is consistent with the fact that more educated individuals experience more variation in lifetime income.

We also identify age, cohort, and time effects in the data for each of the education groups. Regardless of the identification assumptions, we find that time effects are small and mostly homogeneous across groups in East Asia. The effects are much stronger in Latin America, however, where there are large differences across groups, with larger negative effects in downturns and smaller positive effects in upturns for the least educated households. This is especially true under the assumption of no cohort effects, suggesting that less educated individuals in Latin America have more limited capacity to smooth unexpected variations in income and to build up income-earning assets in good times.

Cohort effects, identified under the assumptions on time effects mentioned above, are strong and positive in Taiwan (China) and Thailand for all education groups and strong and positive only for the most educated in Mexico and Peru. Cohort effects among the most educated in Latin America appear to be even stronger than those experienced by their counterparts in East Asia, but overall they are almost flat because of negative or no cohort effects for the other groups.

As with the total population, age effects are estimated with two alternative specifications. When dummy variables for the years are constrained to have zero mean and to be orthogonal to a time trend, the main difference is that in Mexico, Peru, and Thailand age profiles for the least educated and those with secondary education are mostly flat, while they increase monotonically for the most educated. In Taiwan (China) the monotonic increase is observed across all groups. When the age profiles are estimated under the assumption of no cohort effects, the profile for the most educated in Mexico, Peru, and Thailand appears to be more hump shaped than the profiles for those with less schooling. In Taiwan (China) the age profile also becomes much more hump shaped, but there are no differences across education groups. Under some identification assumptions, we thus conclude that the savings behavior of the most educated households in Mexico, Peru, and Thailand is more in line with the life-cycle hypothesis than is the savings behavior of households with less education.

It has been argued that because Latin America is on the verge of a demographic transition similar to that already experienced by East Asia, future demographic trends might bring about an increase in aggregate saving that will reduce the savings gap between the regions. We present simulations indicating that although our life-cycle profiles and cohort effects predict an increase in aggregate savings rates, the changes cannot be attributed to current demographic trends. The increase is driven mainly by strong cohort effects, identified under

the assumption that all the trends in the data can be interpreted as either age or cohort effects. Moreover, although the shape of the age profile is much more in line with the life-cycle model when we use the alternative identifying assumption that there are no cohort effects in savings rates, the estimated hump in the middle of the life cycle is not large enough to generate increases in aggregate savings rates in Latin America large enough to catch up with savings rates in East Asia in the next 20 years. As we emphasize, the simulations should be interpreted with great caution because they are based on very strong identification assumptions and assume that the economic environment in East Asia and Latin America will remain unchanged in the future. This is a rather pessimistic scenario for Mexico and Peru—and perhaps an overly optimistic one for Taiwan (China) and Thailand.

Our results suggest that the projected demographic trends by themselves are unlikely to generate large increases in savings rates under current circumstances. This does not mean that these trends are unimportant. They will play an important role in determining the ability of developing regions to attract capital flows from industrial countries, where capital-labor ratios are projected to be much higher in the future than in developing countries (an issue discussed by Attanasio and Violante 2000).

It is also possible that the estimated age profile of saving, whose shape is responsible for our results, will change as a consequence of the structural changes Latin America is experiencing. Two of these changes are particularly important. The first is changes in labor supply behavior, in particular, an increase in female labor force participation. The second is the shift from public to privately funded pension schemes that has occurred in many Latin American countries in recent years.

Notes

1. The graphs plot the coefficients for the dummy variables for the years that result from putting together two panels with a different mix of countries depending on the region and then estimating fixed effects regressions on each panel, where the dependent variable is the savings rate.

2. The Mexican, Thai, and Taiwanese surveys considered here are strictly comparable in terms of questionnaires, objectives, and sampling techniques. The data for Peru are largely comparable. The 1991 survey excludes some rural areas from the sample, but restricting the comparison to exactly the same geographic areas in all four years does not change any of our conclusions.

3. For Peru we are able to compute only s_1 because of the lack of data by expenditure item. The estimates of s_1 and s_3 for Mexico for 1984, 1989, and 1992 do not coincide exactly with the estimates reported in Székely (1998), although they are produced with the same data. The difference is that Székely measures saving as the difference between disposable income and nondurable consumption but adjusts consumption to include interest payments on debt. We have not made the same adjustment here. As a result, the savings rates are 2.6, 1.8, and 2.3 percentage points higher for s_1 , and 2.4, 1.7, and 2.6 percentage points greater for s_3 , than those reported in Székely (1998). All estimates for Mexico are compatible with those in Attanasio and Székely (1998).

4. Sampling error, induced by small cell sizes, could be interpreted as a time effect. However, it is plausible to assume that sampling error has zero mean and does not exhibit any time trend.

5. Carroll and Summers (1991) interpret such results as a failure of the life-cycle model. Attanasio and Browning (1995) and Attanasio and others (1998) offer reasons, in the context of the United Kingdom and the United States, why the life-cycle model is not necessarily at variance with the data in figure 7.

6. The increase in the last part of the life cycle in Peru is implausibly strong. We tried different specifications for the age polynomials, obtaining similar results. The result is driven by the marked increase in savings rates for the two oldest cohorts.

7. There are two issues. On the one hand, the life-cycle model implies that with a hump-shaped income profile, savings should also be hump shaped, if needs are constant over the life cycle. But as we saw earlier, household size—and therefore needs—change considerably over the life cycle. On the other hand, simple versions of the model imply that saving should decline (and become negative) after retirement. In principle, several data issues might explain why we typically do not observe negative savings rates in the last part of the life cycle. However, the evidence from the four economies seems to indicate that at least for the most educated, saving increases in the last part of the life cycle.

8. Our results for age and cohort effects in Thailand differ from those reported by Paxson (1996). Paxson, who considers the entire population and does not look at education groups, finds an age profile of savings rates that is very flat over the life cycle. While we use slightly different selection criteria (and unlike Paxson do not use expansion factors because of difficulties computing them), the main reason for the difference seems to be that Paxson does not use the data from 1994 and 1996. When we drop those years, we obtain results much closer to hers. The reason is that in the last two years savings rates increase considerably for all cohorts. Our smoothing procedure forces us to interpret these trends as either age or cohort effects, as time effects are assumed to be orthogonal to a time trend. This gives us the rising age profiles for the entire population and for each of the education groups and should remind us of the caveats discussed earlier. Interestingly, our results for Thailand are similar to those that Paxson reports for Taiwan (China).

9. Our results for age and cohort effects in Taiwan (China) are very similar to those reported by Deaton and Paxson (1994b), even though they use data only up to 1990. Our results are also consistent with those reported in Deaton and Paxson (2000).

10. Bloom and Williamson (1999) and Behrman, Duryea, and Székely (1999a) have argued along these lines. Attanasio and Violante (2000) simulate the effects of demographics on private savings and predict a large increase in Latin America as a result of future reductions in the old age dependency ratio. That study, however, focuses on the general equilibrium effects (that is, changes in wage and interest rates).

11. This is a particularly important issue for Mexico, for which 1996, the year we use as a benchmark, is by no means a standard year for comparison. As shown in the previous section, there are strong negative effects from the shock the economy faced in 1995. Assuming that the same conditions will prevail is thus a rather pessimistic scenario. For Taiwan (China) and Thailand the benchmark is 1996, two years before the recent financial crisis. In those economies economic conditions may have changed for the worse, and projections based on the conditions in 1996 can be regarded as rather optimistic.

12. In this exercise the level of savings rates is particularly difficult to pin down because year effects, by definition, are unpredictable.

References

- Attanasio, Orazio P. 1994. "A Cohort Analysis of U.S. Household Saving Behavior." NBER Working Paper 4454. National Bureau of Economic Research, Cambridge, Mass.
- . 1998. "Cohort Analysis of Saving Behavior by U.S. Households." *Journal of Human Resources* 33 (3): 575–609.

- Attanasio, Orazio P., and Martin Browning. 1995. "Consumption of the Life Cycle and over the Business Cycle." *American Economic Review* 85 (5): 1118–37.
- Attanasio, Orazio P., and Henry Hoynes. 2000. "Wealth Accumulation and Differential Mortality." *Journal of Human Resources* 35 (1): 29.
- Attanasio, Orazio P., and Miguel Székely. 1998. "Household Saving and Income Distribution in Mexico." OCE Working Paper Series 390. Inter-American Development Bank, Office of the Chief Economist, Washington, D.C.
- Attanasio, Orazio P., and Giau Lucca Violante. 2000. "The Demographic Transition in Closed and Open Economies: A Tale of Two Regions." RES Working Paper Series 412. Inter-American Development Bank, Research Department, Washington, D.C.
- Attanasio, Orazio P., James Banks, Costas Meghir, and Guillermo Weber. 1998. "Humps and Bumps in Lifetime Consumption." *Journal of Business and Economic Statistics* 17 (1): 22–35.
- Barro, Robert J., and Jong-Wha Lee. 1993. "International Comparison of Educational Attainment." *Journal of Monetary Economics* 32 (3): 363–94.
- Behrman, Jere, Suzanne Duryea, and Miguel Székely. 1999a. "Aging and Economic Opportunities" OCE Working Paper Series 405. Inter-American Development Bank, Research Department, Washington, D.C.
- . 1999b. "Schooling Investments and Macroeconomic Conditions: A Micro-Macro Investigation for Latin America and the Caribbean." OCE Working Paper Series 407. Inter-American Development Bank, Research Department, Washington, D.C.
- Bloom, David E., and Jeffrey G. Williamson. 1999. "Demographic Transitions and Economic Miracles in Emerging Asia." *World Bank Economic Review* 12 (3): 419–56.
- Browning, Martin, Angus Deaton, and M. Irish. 1985. "A Profitable Approach to Labor Supply and Commodity Demands over the Life Cycle." *Econometrica* 53 (3): 503–43.
- Carroll, Chris, and Lawrence H. Summers. 1991. "Consumption Growth Parallels Income Growth: Some New Evidence." In B. D. Bernheim and J. B. Shoven, eds., *National Saving and Economic Performance*. Chicago: University of Chicago Press.
- Deaton, Angus, and Christina Paxson. 1994a. "Intertemporal Choice." *Journal of Political Economy* 102 (3): 437–67.
- . 1994b. "Saving, Growth, and Aging in Taiwan." In D. Wise, ed., *Studies in the Economics of Aging*. Chicago: University of Chicago Press.
- . 2000. "Growth and Saving among Individuals and Households." *Review of Economics and Statistics* 82 (2): 213–25.
- Heckman, James, and Richard Robb. 1987. "Using Longitudinal Data to Estimate Age, Period, and Cohorts Effects in Earnings Equations." University of Chicago, Department of Economics.
- Lora, Eduardo, and Carmen Pagés. 2000. "Hacia un envejecimiento responsable: Las reformas de los sistemas de pensiones en America Latina." Inter-American Development Bank, Research Department, Washington, D.C.
- MaCurdy, Tom, and Thomas Mroz. 1995. "Measuring Macroeconomic Shifts in Wages from Cohort Specifications." Stanford University, Department of Economics, Stanford, Calif.
- Paxson, Christina. 1996. "Savings and Growth: Evidence from Micro Data." *European Economic Review* 40 (2): 255–88.
- Székely, Miguel 1998. "Nivel y distribución del ahorro de los hogares en México." *El Trimestre Económico* 65 (2): 263–314.
- United Nations. 1998. *World Population Prospects 1950–2050: The 1998 Revision* (Database). United Nations New York: Population Division.
- World Bank. 1999. *World Development Indicators 1999*. Washington, D.C.

Comment on “Household Saving in East Asia and Latin America: Inequality, Demographics, and All That,” by Orazio P. Attanasio and Miguel Székely

Robert Holzmann

In economics we are familiar with the concept of lenders of last resort—the central bank for the financial system, the government for social protection programs. At conferences we are familiar with commentators of last resort—people already on the panel. I took on the role of commentator because of the first paragraph of Orazio P. Attanasio and Miguel Székely’s article, which emphasizes the importance of savings for smoothing unexpected variations in the income of households and the resulting welfare implications in view of missing credit and insurance markets. Emphasizing the ability to transfer income between periods—to save—and thus to avoid costly coping mechanisms that interrupt human capital accumulation is very much in line with the social risk management framework that the Social Protection Sector of the World Bank has developed for its strategy paper (Holzmann and Jorgensen 2000; World Bank 2001). It is also at the core of the Bank’s *World Development Report 2000/2001* (2000).

Measured against this motivation and objective, the article fares reasonably well, but more needs to be done. It is an interesting and rich article but is still very much a work in progress—much of the information at hand seems not yet distilled, and the work relies on graphical presentation rather than econometric analysis.

Moreover, the article wants to do too much. It might be useful to think about splitting it into two articles—one dealing with the microeconomic savings profile, the testing of the life-cycle hypothesis, the welfare implications, and so forth, and one addressing the macroeconomic issues, such as projections of future national savings rates, the implications for financial flows between countries and regions, and so forth.

Finally, and perhaps most important, the article currently lacks analysis and conclusions that link savings profiles at different education levels with the motivating question of the first paragraph—can and do the poor save and thus do they have the capacity to ride out income fluctuations? In view of the provisional character of the

Robert Holzmann is director of the Social Protection Sector at the World Bank.
Annual World Bank Conference on Development Economics 2000
©2001 The International Bank for Reconstruction and Development / THE WORLD BANK

article, my comments are a selection of possibly arbitrary questions, comments, and suggestions.

Selection of Countries

The article attempts to compare Latin America and East Asia, whose aggregate national savings rates differ significantly. This is an important objective, since much of the difference in economic development between the two regions has been related to differences in savings behavior. But the choice of economies for the comparison—Mexico, Peru, Taiwan (China), and Thailand—might not achieve the objective, because it means contrasting two Latin American countries with average performance with two (perhaps former) Asian tigers. To achieve the objective, it would be better to compare laggards and leaders within each region, assuming, of course, that comparable data for such comparisons are available, which may not be the case.

Discrepancies between Macroeconomic and Aggregated Savings Data

Table 1 compares the aggregate private savings rate obtained from macroeconomic accounts with the household savings rate obtained from aggregated cross-section savings data for households. While the two measures are similar for Thailand (35.9 and 29.7 percent of GDP), they are very different for Mexico (25.4 and 9.5 percent of GDP), Peru (19.0 and 9.6 percent of GDP) and Taiwan (26.8 and 49.1 percent of GDP). Unless such differences can be convincingly explained (by public or corporate saving, for example)—and in this paper they are not—they cast doubt on the quality of the household cross-section data.

Do the Poor Save?

The article does not address the issue of whether the poor save, and the savings profiles do not provide a clear answer. Some savings profiles for the least educated group (as a proxy for the lowest income group) suggest that in Mexico the poor may have the capacity to save. Other profiles suggest they do not. Figure 15 (with no cohort effect) suggests a hump-shaped savings profile; figure 16 suggests dissaving around 1995, the year of the economic crisis, while such an effect is only slightly visible in figure 14 (with the assumption of a cohort effect). The opposite is suggested for Thailand, where the picture is reversed (figure 16), but where a sharp drop in dissaving in some years is suggested when a cohort effect is taken into account (figure 14). I hope future versions of the article will shed some light on this issue.

Linking Savings Data Analysis with Institutional Information

The capacity and willingness to save are determined by institutional characteristics, such as coverage by a public pension system and access to free basic education and

health services. The current analysis does not use institutional differences to explain differences in savings profiles across countries. The graphical presentation uses only one concept of saving that approximates financial saving: disposable income minus expenditure (including expenditures on housing, health, and education). Other available savings data, based on different concepts of saving, exclude one or the other investment-type expenditure, but such data are not discussed. Linking institutional differences to differences among alternative savings definitions might help clarify some of the savings puzzles. Similarly, a more comprehensive discussion of the life-cycle profile might help explain the existence or lack of public pension schemes. In Thailand a mandatory scheme for the formal sector was introduced only in 1999, while in Mexico public pensions were in place during the entire period investigated (Social Protection Pensions Primer Papers).

References

- Holzmann, Robert, and Steen Jorgensen. 2000. "Social Risk Management: A New Conceptual Framework for Social Protection, and Beyond." *Social Protection Discussion Paper 0006*. World Bank, Washington, D.C.
- Social Protection Pensions Primer Papers. Various authors and years. [<http://www.worldbank.org/pensions>].
- World Bank. 2000. *World Development Report 2000/2001: Attacking Poverty*. New York: Oxford University Press. [www.worldbank.org/poverty/wdrpoverty].
- . 2001. "Social Protection Sector Strategy: From Safety Nets to Spring Board." World Bank, Social Protection Sector, Washington, D.C. [<http://www.worldbank.org/sp>].

THEMES AND PARTICIPANTS FOR THE 13TH ANNUAL WORLD BANK CONFERENCE ON DEVELOPMENT ECONOMICS (ABCDE)

MAY 1–2, 2001

Globalization and Inequality

Globalization and Inequality: Historical Trends
Fear of Globalization: The Human Capital Nexus
Growth and Inequality: The Role of Foreign Trade and Investment
Geography and International Inequalities: The Impact of New Technologies

Health and Development

The Role of Nongovernmental Organizations in Providing Health Care
A Patent Policy Proposal for Global Diseases
Health, Income, and Economic Development
Human Capital, Longevity, and Economic Growth: A Quantitative Assessment
of Full Income Measures

Sudhir Anand • Patrick Asea • Abdul Bayes • Anne Case
Daniel Cohen • Richard N. Cooper • Rachel Gesami • Jean O. Lanjouw
Juan Luis Londoño • Kevin H. O'Rourke • Siddiqur Rahman Osmani
Tomas Philipson • Andrés Rodríguez • Morten Rostrup • Rodrigo Soares
Anthony J. Venables • Beatriz Zurita



THE WORLD BANK

1818 H Street N.W.
Washington, D.C. 20433, U.S.A.
Telephone: 202-477-1234
Facsimile: 202-477-6391
Internet: www.worldbank.org
E-mail: feedback@worldbank.org

The Annual World Bank Conference on Development Economics (ABCDE) brings together the world's leading scholars and development practitioners for a lively debate on state-of-the-art thinking in development policy and the implications for the global economy. The twelfth conference dealt with four difficult topics in development: new development thinking, with Paul Collier, Dani Rodrik, Jan Willem Gunning, and Karla Hoff; crises and recovery, with William Easterly, Roumeen Islam, and Joseph E. Stiglitz; Ricardo J. Caballero and Mohamad L. Hammour; Eisuke Sakakibara; and Daniel Lederman, Ana María Menéndez, Guillermo Perry, and Joseph E. Stiglitz; corporate governance and restructuring—lessons from transition and crises, with I.J. Alexander Dyck and Gérard Roland; and social security, public and private savings, with J. Michael Orszag and Peter R. Orszag and Orazio P. Attanasio and Miguel Székely.

ALSO IN THIS VOLUME:

COMMENTS BY

Abhijit Banerjee • Andrew Berg • Nancy Birdsall • François J. Bourguignon • John S. Earle • Robert Holzmann • Fukunari Kimura • Rafael La Porta • Gustav Ranis • Charles C. Soludo • Alejandro Werner



COVER DESIGN BY COMMUNICATIONS DEVELOPMENT INCORPORATED

ISBN 0-8213-4981-3