Lessons from East Asia and the Global Financial Crisis

Edited by
Justin Yifu Lin and Boris Pleskovic
Lessons from East Asia and the Global Financial Crisis
ANNUAL WORLD BANK CONFERENCE ON DEVELOPMENT ECONOMICS

Previous titles in this series:

ABCDE 2009, Global: People, Politics, and Globalization
ABCDE 2008, Global: Private Sector and Development
ABCDE 2008, Regional: Higher Education and Development
ABCDE 2007, Global: Rethinking Infrastructure for Development
ABCDE 2007, Regional: Beyond Transition
ABCDE 2006, Europe: Securing Development in an Unstable World
ABCDE 2006: Growth and Integration
ABCDE 2005, Europe: Are We on Track to Achieve the Millennium Development Goals?
ABCDE 2005: Lessons of Experience
ABCDE 2004, Europe: Economic Integration and Social Responsibility
ABCDE 2004: Accelerating Development
ABCDE 2003, Europe: Toward Pro-Poor Policies—Aid, Institutions, and Globalization
ABCDE 2003: The New Reform Agenda
Lessons from East Asia and the Global Financial Crisis

Edited by
Justin Yifu Lin
and Boris Pleskovic

THE WORLD BANK
Washington, D.C.
# Contents

ABOUT THIS BOOK ix

INTRODUCTION 1

Justin Yifu Lin and Boris Pleskovic

OPENING ADDRESS 15

Jeung-Hyun Yoon

OPENING ADDRESS 19

Learning from the Past to Reinvent the Future

Justin Yifu Lin

KEYNOTE ADDRESS 31

The Global Financial Crisis: Causes and Policy Responses

Il SaKong

KEYNOTE ADDRESS 37

Development Prospects in Light of the Global Financial Crisis

Anne O. Krueger

KEYNOTE ADDRESS 47

The Global Crisis: Is It Over Yet?

Simon Johnson

Industrial Policy and Development 61

Industrial Policy and Development: A Political Economy Perspective

James A. Robinson

COMMENT 80

Bert Hofman
Industrial Policy: Can We Go Beyond an Unproductive Confrontation? 83
Ha-Joon Chang

COMMENT 110
Karen R. Polenske

Social Capital, Institutions, and Development
A Matter of Trust: Social Capital and Economic Development 119
Partha Dasgupta

“Individual” Social Capital, “Social” Networks, and Their Linkages to Economic Game 157
Masahiko Aoki

COMMENT 174
Mariano Tommasi

Financial Crisis and Regulation
Reform of the Global Regulatory System: Perspectives of East Asia’s Emerging Economies 187
Yung Chul Park

Financial Crisis and the Paradox of Underregulation and Overregulation 213
Joshua Aizenman

COMMENT 235
Jong-Wha Lee

The Road Ahead to a Sustainable Global Economic System
Lessons from the Recent Financial Crisis for Reforming National and International Financial Systems: The Road Ahead to a Sustainable Global Economic System 241
Stijn Claessens

COMMENT 275
Yoon Je Cho

A Sustainable Global Economic System after the “Great Recession”? Some Lessons from History 281
Giovanni Zanalda

COMMENT 301
Jeromin Zettelmeyer
Innovation and Competition

Innovation and Financial Globalization 309
Philip R. Lane

Innovation, Competitiveness, and Growth: Korean Experiences 333
Sungchul Chung

COMMENT 358
Jean-Jacques Dethier

CLOSING SUMMARY 367
Alan Gelb and Shahrokh Fardoust

CLOSING REMARKS 373
Young Geol Lee
About This Book

The Annual World Bank Conference on Development Economics (ABCDE) is a forum for discussion and debate of important policy issues facing developing countries. The conferences emphasize the contribution that empirical 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, to reflect the comments made by discussants or from the floor, while most 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. Unless otherwise noted, participants’ affiliations identified in this volume are as of the time of the conference, June 22–24, 2009.

The planning and organization of the June 2009 conference was a joint effort by the Government of the Republic of Korea, the Korea Development Institute (KDI), and the World Bank. We gratefully acknowledge timely and valuable contributions made by all the members of the Steering Committee and several anonymous reviewers.

We would also like to thank Alan Gelb and Aehyung Kim for their insightful advice and Leita Jones, conference organizer, whose excellent organizational skills helped to ensure a successful conference.

Finally, we thank the editorial staff from the World Bank’s Office of the Publisher for their efforts in pulling this volume together.
The Annual Bank Conference on Development Economics (ABCDE) brings together leading policy makers, academics, and researchers to advance the debate on key problems on development. The 2010 ABCDE was held in Seoul, Republic of Korea (June 22–24, 2009), in the midst of the deepest and longest world recession since the Great Depression.

This year’s ABCDE included sessions on the following themes: industrial policy and development; social capital, institutions, and development; financial crisis and regulation; the road to a sustainable global economic system; and innovation and competition. In light of the global financial crisis, speakers touched on fundamental questions: What caused the current crisis, and how can the world economy recover? Are the standard prescriptions of development economics adequate to the task? Should developing countries alter their basic growth strategies? What is the proper role of the state? Should developing countries reexamine their commitment to free trade? How can global imbalances be rectified (especially between China and the United States)? Within the globalized financial system, how can regulation be improved? In attempting to answer these questions, many of the speakers searched for solutions in the lessons offered by the experience of Korea and other East Asian countries, which reacted with varying degrees of success to the financial crisis of the late 1990s.

This volume includes selected papers from the conference as well as keynote addresses by Il SaKong, chairman of the Korean G-20 Summit Coordinating Committee, and two distinguished economists: Anne Krueger, Stanford University and Johns Hopkins University, and Simon Johnson, Massachusetts Institute of Technology.
In their opening remarks Jeung-Hyun Yoon, minister of strategy and finance of the Republic of Korea, and Justin Yifu Lin, senior vice president and chief economist of the World Bank, emphasize the development experience of East Asia, especially Korea’s rapid growth over the past half century.

Yoon underscores that the global economy is still in the midst of crisis and that at this historic moment it is imperative to set a new direction for the world economy. Yoon recalls a crucial lesson of the Great Depression: international coordination is of the highest importance in overcoming a global financial crisis. He states that the two recent G-20 summits were a turning point in facing the challenges posed by the current crisis, and next year, when Korea assumes the chair of the G-20, his country will continue to solidify the G-20 framework.

Yoon maintains that to sustain long-term growth—within the context of rising population, climate change, protectionism, and poverty—“green growth” must become the engine of global growth. At the same time, in spite of the crisis, he emphasizes that we must not give up the grand principle of free trade and retreat into protectionism.

Yoon adds that, regardless of the urgency of rescuing the global financial system, the industrial countries must not lose sight of their obligation to reduce poverty and income inequality around the world. He believes that Korea, given its rapid rise from profound poverty to advanced development, has a special role to play as a bridge between the developing and advanced countries. In addition, he states that Korea will expand its knowledge-sharing programs and triple official development assistance by 2015.

Lin notes some signs of recovery, but also many troubling indications that the crisis is far from over. In particular, he observes that unemployment is likely to remain a severe problem for quite some time, especially in developing countries. The crisis, he says, has shaken confidence in the general premise that markets deliver socially superior outcomes, and he encourages economists to take this as an opportunity to rethink the fundamentals of economic development. Lin states that some of the main themes of the ABCDE—the role of institutions and the boundaries between markets and the state—will be at the center of the debate in development economics for years to come.

Turning to the economic history of the East Asian countries, Lin draws a few general lessons for development as well as lessons for the crisis. He notes that East Asia was at the same level of development as Africa and South Asia in the 1950s, but since then has grown far more rapidly than any other region in the world. Korea; Taiwan, China; Hong Kong, China; and Singapore have maintained an annual growth rate of 10 percent over two or three decades—an unprecedented explosion of economic growth. Market mechanisms, according to Lin, were the foundation of this long period of sustained rapid growth, but government has also played an active role in the economy. Lin observes that these countries did not simply follow the prescriptions of standard development theory, but they did follow strategies that were consistent with their comparative advantage and they let the market allocate resources. Lin concludes that the experience of East Asia suggests the need to formulate a new approach to
development, with the government facilitating industrial progress through information, infrastructure, and incubation of new industries.

The East Asian financial crisis of the late 1990s provides a number of broadly relevant lessons, but the most valuable may be that carefully crafted unorthodox policies may work better than the standard prescriptions of development economics. Lin observes that during its “lost decade” of the 1990s, Japan enacted conventional measures, while China made major investments in transportation infrastructure in line with the concept of “bottleneck-releasing projects.” China’s plan proved much more successful, and its success, according to Lin, shows not only that a well-designed stimulus can enhance growth but also that this type of growth-enhancing, bottleneck-releasing investment can lead to sustainable growth, higher fiscal revenue, and lower debt. Lin concludes that the best path for developing countries may be to release bottlenecks, while in developed countries it may be to stimulate the “green economy” and address climate change, as Korea has pledged to do.

Keynote Addresses

Il SaKong, chairman of Korea’s G-20 Summit Coordinating Committee, delivered the first keynote address. SaKong outlines the major causes of the crisis, the global community’s response, and the role that Korea may play. He calls for reform of the Bretton Woods institutions and suggests agenda items for the next G-20 summit, to be held in Pittsburgh in September 2009.

Many factors contributed to the global financial crisis, he says, but clearly it began in the United States. The U.S. government took the initial steps to deal with the crisis, but when it became clear that the crisis would require a global response, the leaders of 20 nations convened in Washington to coordinate that response. SaKong maintains that this meeting reflected the historical shift in economic power that had taken place in the previous two decades, and he identifies the G-20 meeting as an event of historical importance for global governance.

At that meeting and a subsequent summit in London, G-20 leaders agreed upon bold fiscal measures to stimulate demand and to resist protectionism. According to SaKong, Korea played a special role with respect to emerging and developing countries in two ways: (1) having known within living memory severe poverty and recognizing the importance of trade for development, Korea worked as an intermediary between the developed and developing countries; and (2) Korea, along with the United Kingdom, emphasized concrete deliverables and brought in the International Monetary Fund (IMF) and the World Trade Organization (WTO).

The next G-20 summit will be held in Pittsburgh. It is too early to talk of exit strategies, SaKong warns: declaring victory prematurely, as the United States did during the Great Depression and Japan did in the 1990s, could make the crisis much worse. SaKong emphasizes the importance of following through on the London agreements and devoting special attention to the developing and emerging economies. Finally, SaKong affirms that the work of the international financial institutions must be
improved and the voices of emerging economies strengthened. He notes that eight of 24 seats on the IMF's board are held by Europe, which does not correspond to Europe’s economic weight today.

Anne Krueger asserts that the efficiency of the financial system is a critical determinant of economic growth. While the international financial system and regulatory framework need to be strengthened, she warns that to fundamentally alter development policy would be wrong. In making her argument, Krueger examines the role of the financial system, the outlook for the global economy, how different countries have been affected by the crisis, and lessons for the future.

Historically, the financial system has played a vital positive function alongside the real economy, and Krueger emphasizes that this should not be forgotten when pondering how the international financial architecture might be modified. Nor should we forget, according to Krueger, that in the long term, countries that have experienced financial crises have grown more rapidly than those that have not.

Krueger reviews the causes of the current crisis and notes that low interest rates certainly contributed to the problems of the U.S. housing sector. She argues, however, that interest rates were low in part as a result of global imbalances, especially the imbalance between U.S. current account deficits and Chinese surpluses. In addition to stronger financial regulation, Krueger insists that the world also needs to create a system to prevent or mitigate unsustainable global imbalances.

Krueger notes that the crisis has turned out to be much broader and deeper than expected, yet the effects on developing countries, while very serious, have been far less serious than they might have been. Many developing countries were in much better shape to face this crisis than previous ones. Many had grown rapidly in the years before the crisis and in many cases had established effective fiscal policies. As a rule, Krueger observes, the larger the initial fiscal deficits and debt and the more dependent the countries on exports of primary commodities, the greater has been the impact of the current economic slowdown. One of the lessons for the future, Krueger says, is that developing countries should adopt structural fiscal rules to minimize the impact of global slowdowns.

Turning to the global outlook, Krueger forecasts slower growth than in recent years, with higher real interest rates and possibly smaller global imbalances. Krueger notes that public debt in the industrial countries is increasing rapidly and the aging of the population will put increasing pressure on government spending. The result, she says, will be higher interest rates and scarcer capital in developing countries. This situation will favor countries that have abundant unskilled labor and those that have invested in human capital.

Krueger observes that no developing country has sustained a successful long-term development strategy without opening to international markets. On that basis, she concludes that strengthening the open multilateral trading system should remain a priority. Krueger also advocates having more developing countries formally enter the WTO: too many are free riders. During the crisis, countries that had not formally joined the WTO were able to raise tariffs. Since it is easier to raise tariffs than to lower them, Krueger expects this step to damage these countries in the
long term. As the Korean experience shows, the opportunity to use international markets is vitally important for rapid economic growth, especially in the early stages of development.

Simon Johnson states that, while confidence is returning to the financial markets, the underlying problems remain and indeed are getting worse. Johnson maintains that the gargantuan financial sectors of the United States and Western Europe have endangered the entire global economic system and that the U.S. government’s financial injections into the U.S. banking system are making the global system even more vulnerable.

According to Johnson, the top U.S. economic officials in the Obama administration portray the current crisis as a “once in a century” event that requires a massive macroeconomic policy response, but only a minimal regulatory response. Johnson says that this response reduces the chances of bank failure and thereby stabilizes the sector, but it represents an enormous, unconditional subsidy for the banking system.

Johnson maintains that a major political and economic structural change has been under way in the United States since the 1980s, and this change was mirrored, even exaggerated, in Western Europe. Johnson puts this into a larger historical context, observing generally that powerful groups inevitably rise up during long economic booms. These groups, he says, accumulate disproportionate political power, which they use for their own benefit and often to the detriment of society as a whole. Johnson cites historical cycles of deregulation and reregulation in the United States, such as Andrew Jackson’s battle against big banks in the 1830s, Teddy Roosevelt’s “busting” of the big railroad trusts at the end of the nineteenth century, and financial reorganization following the crash of 1929.

In today’s crisis situation, the big American banks, according to Johnson, are able to extract enormous subsidies from the U.S. government. Some might call this phenomenon regulatory capture, but Johnson says it might be more accurate to call it “state capture” or “oligarchy.” Indeed, it might be termed “intellectual capture” inasmuch as the powerful U.S. financial sector over the past three decades has persuaded people (including those serving in the Senate and House of Representatives) that unfettered finance benefits the economy as a whole.

According to Johnson, at the same time the U.S. financial sector was accumulating inordinate political power and undermining political institutions during the last few decades, emerging markets opened to capital flows on a much larger scale than ever before. This has turned the crisis in the U.S. and European financial sectors into a worldwide crisis.

Johnson asks, what, if anything, can break a crisis that has been built upon the concentration of political and economic power? The obvious answer is the bankruptcy of oligarchic enterprises. This solution, Johnson says, is easy in economic terms, but very difficult in political terms. Such solutions come about fairly often in developing countries, but in the U.S. context, the situation is entirely different: resources are sufficient for a bailout and the IMF will not be called in. He concludes that in the United States the big bankers have won for now, but they are unlikely to win in the long run. Due to their lack of restraint, Johnson argues, bankers will
inevitably overreach to a point that ignites popular outrage. Johnson is confident that ultimately this will lead to a breakup of the banks so that they are no longer “too big to fail.”

**Industrial Policy and Development**

*James Robinson* argues on the basis of both theory and empirical evidence that industrial policy can, at least under certain circumstances, stimulate growth and development. But in actual practice, he says, more often than not industrial policy seems to impede growth. Industrial policy has succeeded in a few places, most notably in Korea and Taiwan, China, but in other places, especially Latin America and Sub-Saharan Africa, it has failed spectacularly. Robinson contends that we are in sore need of a positive theory of industrial policy to explain why industrial policy has worked in some countries but failed in others, and he begins to outline the framework of such a theory.

Robinson argues that the success or failure of industrial policy in any given country depends on the political equilibrium in that country. To develop a positive theory of industrial policy, Robinson says, we need to analyze such policy as an equilibrium outcome, using the methods of political economy. Robinson contends that industrial development is not a question of proposing good policies—economists have been doing that for decades. Instead, he argues, industrialization is promoted (or hindered) by the political choices made by any given society, and it is promoted if and only if it is aligned with the interests and institutions of those in power. Robinson insists that economists and international institutions have to take this fundamental fact into consideration when trying to promote industrialization and either formulate policy in such a way that it provides incentives for those in power or alter the political equilibrium in such a way that it favors industrialization.

Reviewing the cases of Ghana, Zambia, and other countries, Robinson argues that the most important differences between the cases of unsuccessful and successful industrial policy are political rather than economic or administrative. Political decisions, he says, in many cases lead to uneconomic allocations of capital investment. He presents a series of examples from Africa and Latin America and concludes that allocation in those cases was fundamentally driven by political criteria with little regard for economic criteria. He emphasizes that what was different in the East Asian “miracle” economies was that the state bureaucracy was allowed to develop policies based more on rational economics than on political motivations. Ultimately, he says, the difference really lies in the institutions implementing the policy: it is the distribution of political power in society—both de jure (institutional) and de facto—that determines the choice of institutions and policies.

Robinson argues further that those with power seek not only to maximize their income today but also to maintain their power. For this reason, Robinson says, even policies that would increase the incomes of today’s elite will not be chosen if they undermine the elite’s grip on power. Those who stand to benefit only indirectly from industrialization may not have a sufficient incentive to advance it, since it inevitably brings social change and undermines the political status quo.
Based on his review of political economy, Robinson concludes that economists and policy makers must change the way they think about industrial policy. To give advice that would foster industry, one has to understand the political equilibrium at work within a country and either change it or work within it. The successful industrial policy of East Asian countries, he says, reflects the very different political equilibrium that emerged in that part of the world.

Ha-Joon Chang takes up industrial policy and attempts to lay out the positions of both its opponents and proponents, highlighting the areas in which there is substantial agreement and suggesting ways to advance the debate constructively.

He begins with a brief review of the history of debate on industrial policy, looking more broadly at historical experience than is usual, so as to provide a larger framework for understanding industrial policy. He takes the broad view in order to establish a minimum common empirical understanding: specifically that industrial policy can work, sometimes spectacularly well, although it can also fail, sometimes spectacularly.

Chang begins his review of the general debate in the 1970s, following the rise of Japan. But he also looks further back in history and points out other examples of apparently successful industrial policy: postwar France, Finland, Norway, Austria; Britain in the eighteenth and nineteenth centuries; the United States, Germany, and Sweden in the late nineteenth and early twentieth centuries; and so on. Chang concludes from these cases, and the more recent East Asian cases, that industrial policy has been at least partly responsible for the remarkable economic growth that took place in these countries. He further states that regardless of whether there is absolute proof that industrial policy contributes decisively to growth, it is still possible to discuss constructively how to strengthen its positive effects.

In that spirit, in the main part of his paper, Chang critically examines some of the key issues in the debate on industrial policy. He first looks at two issues commonly brought up on a theoretical basis—targeting and the ability of government bureaucrats to “beat the market.” In regard to targeting, Chang argues that some degree of targeting is an inevitable part of industrial policy. In regard to whether the state can “beat the market,” Chang enumerates several examples in which governments made investment decisions that seemed to run counter to the market and yet proved to be some of the most successful investments in history. Chang does not believe that the government officials who made these investments knew more or were smarter than businessmen; it is more likely, he says, that government officials were in a position to look at things from a longer-term national point of view, which proved advantageous in certain cases.

In looking at political economy, Chang identifies the main factor leading to success as the government’s willingness and ability to impose discipline on the industries it supports. He also asserts generally that good export performance is critically important to economic development. Chang observes that the advocates of industrial policy often do not fully appreciate how critical export performance is, while many opponents do not fully appreciate that export success also requires industrial policy.

In conclusion, Chang urges adversaries in the debate on industrial policy to take a more pragmatic approach and focus on practical issues and the vast middle ground that they share. He cautions policy makers and development economists not to let the
lack of consensus on industrial policy be an excuse for inaction. Many success stories, he says, were based on measures that were not perfect, but were “good enough.”

**Social Capital, Institutions, and Development**

The conferees in the session on social capital, institutions, and development agree that the concept of “social capital” has a powerful intuitive appeal and has come to be seen as a crucial economic factor, yet it is extremely difficult to measure and there is not a precise understanding of how and to what extent it leads to economic progress. The papers presented in the session take a game-theoretic approach to examine the concept in an attempt to advance understanding of how social capital works.

*Partha Dasgupta* observes that scholars have looked at social capital in different ways, but they all place social capital in the space between the individual and the state. Dasgupta argues that the concept of social capital should be understood as interpersonal networks, where members develop and maintain trust in one another to keep their promises by the device of “mutual enforcement” of agreements. He observes that it is not the interpersonal networks themselves that lead to economic progress, but, at a more fundamental level, the trust among people who act within networks. Dasgupta maintains that social capital is a means of creating trust—the underpinning of interpersonal networks and impersonal institutions alike.

Dasgupta analyzes trust through a series of thought-experiments applying the techniques of game theory. In his analysis, he classifies four types of social environments in which the promises people make to one another are credible: mutual affection (exemplified by the household), pro-social disposition (exemplified by common understandings of citizenship or personal integrity), external enforcement (exemplified by the rule of law), and mutual enforcement (exemplified by social norms or Nash equilibrium rules of behavior). In examining the last of these four classes of social environments, Dasgupta finds that social norms are able to sustain trust and cooperation only so long as people have reason to value the future benefits of cooperation; however, he goes on to describe conditions under which mutual enforcement breaks down, in some cases even where there is no “real” change, only a change in certain beliefs held in society.

Dasgupta examines also the workings of interpersonal networks, whether inherited or created, to examine how they contribute to economic well-being. He makes some general observations about social networks—for example, that they take effort to maintain or create but the cost declines the more they are used and trust grows. He also highlights the findings of other scholars—for example, that membership in social networks is an important component of human capital and that even weak network ties are important and beneficial.

Completing his analysis, Dasgupta describes various externalities arising from social networks as well as ways in which social networks complement markets and ways in which they can work in opposition to markets.
In an interesting appendix, Dasgupta argues that the best way to try to quantify social capital is through examination of total factor productivity, and he demonstrates how an increase in trust among people results in an increase in the economy’s wealth.

Masahiko Aoki examines the concept of “social capital” and how it might be useful in understanding economic performance. In attempting to answer these questions, Aoki posits a social-exchange game as the central theoretical concept of his analysis and describes social capital as an individual investment owned by individuals. He observes that social exchange shares many of the characteristics of economic exchange; however, unlike economic exchange, social exchange can be performed without an explicit agreement: unspecified obligations of reciprocity suffice. Aoki’s line of reasoning leads him to define social capital as “the present value sum of the agent’s expected social payoffs over time.” He sees an individual’s social capital as the object of individual investment, but notes that it depends not only on the individual’s actions but also on a system of beliefs regarding one another’s actions. He examines how individuals invest in their own social capital and theorizes how certain social networks or patterns of behavior arise from that.

But how are noneconomic social-exchange games linked to economic games, and how is social capital related to economic outcomes? Aoki observes that economic and noneconomic games may share players (individuals and groups), and norms may evolve through interactions within the two games. Thus individual social capital may be directly relevant to the understanding of economic performance. Aoki observes that the social capital of each agent in a social-exchange game can be used as incentive for cooperative behavior in the economic games. Thus he argues that economic and social-exchange games are often linked and that social beliefs do not evolve independently of economics. Aoki gives illustrations from distant history as well as contemporary Wall Street and modern high-tech industries. He maintains that corporate social capital may contribute to the prospects of long-term profitability.

Financial Crisis and Regulation

Yung Chul Park analyzes the causes and consequences of the global financial crisis, particularly with respect to reserve currency liquidity in East Asia, and makes suggestions for the reform agenda. Park traces with concern the growth of maturity mismatches and currency mismatches in East Asian economies. These twin mismatches pose a threat to the financial sector and ultimately could trigger a currency crisis.

Park observes that capital flows into East Asia have dropped precipitously as a result of the global financial crisis. Since East Asian countries use dollars and euros as reserve currencies, the crisis has squeezed reserve currency liquidity drastically. East Asian countries do not have sufficient foreign assets to alleviate the crunch.

A major cause of the liquidity crisis, according to Park, has been the mismatch between the maturities of foreign assets and liabilities on bank balance sheets. Park states that mismatches of this sort cannot be avoided entirely, since they arise from
normal banking operations. But Park says that such mismatches are dangerous, since they can lead to banks increasingly financing local currency loans with foreign currency. Thus the twin mismatches can easily lead to depreciation of the local currency. This, in turn, exacerbates the problem of currency mismatching and can lead to a currency crisis.

Park contends that currency mismatches were a prominent feature of the Asian crisis of 1997–98. They are not as prominent—at least so far—in the current crisis, but the current crisis is far from over. Although East Asian governments put in place domestic national regulation following the 1997–98 crisis, none of the measures is entirely reliable, and East Asia has not been shielded in this case. So far, according to Park, the danger of a currency crisis is limited to Korea, Indonesia, and Singapore. But he cites other economists who have identified maturity mismatch as one of the main sources of financial instability in the current crisis.

Park observes, as others at the conference have, that global financial integration has brought additional instability to emerging economies, and he contends that what is really needed is a set of new global financial institutions to provide effective supervision and regulation and serve as a global lender of last resort. In addition, he says, more prudent regulation of capital movements in emerging countries is needed.

Looking at past efforts to improve global financial regulation, however, Park does not see reason to hope that a global system will come together quickly, if at all. East Asia, he says, would be wiser to look at regional regulatory collaboration as the second-best and more realistic option. There is already a regional liquidity support system called the Chiang Mai Initiative Multilateralization, which might be the nucleus of a future regional regulatory and supervisory coordinator.

Joshua Aizenman notes two opposite tendencies driving the financial regulatory cycle: underregulation when times are good, and overregulation as a reaction to crisis. Aizenman describes the cycle or paradox of financial regulation this way: effective regulation leads to complacency, reducing demand for financial regulation; reduced demand leads to underregulation and eventually crisis; financial crisis leads to overregulation. This, according to Aizenman, is precisely what happened in the United States and the Organisation for Economic Co-operation and Development countries in the run-up to the present crisis.

Aizenman claims that asymmetric information is at the root of this paradox: crises that have been avoided are imperceptible and hence not part of the political discourse; projects and entrepreneurs who do not receive financing during periods of overregulation are similarly invisible and not part of the political discourse.

Aizenman presents a model showing the imbalance between individual demand for regulation and the socially optimal level. In conclusion, Aizenman outlines a regulatory structure that would mitigate the above-mentioned concerns. He advocates above all greater centralization and transparency, arguing that this would mitigate the problems of asymmetric information. With better, more up-to-date information, regulators, he believes, would be able to monitor and assess systemic risk in real time. He argues in addition that greater regulatory independence would reduce the ten-
dency to underregulate in good times, as would adoption of global standards of minimum prudential regulations and information disclosure.

The Road Ahead to a Sustainable Global Economic System

*Stijn Claessens* reviews the causes of the current global financial crisis—those that are familiar from previous crises as well as those peculiar to this crisis: the use of complicated, opaque financial instruments, integration of financial markets both nationally and internationally, highly leveraged financial institutions, and the prominent role of overextended households, especially in the United States.

The global crisis has evoked large-scale government intervention, which, Claessens states, appears to have been effective in stabilizing financial systems and restoring basic confidence. But Claessens observes that massive government intervention also introduces distortions into the market—direct distortions, such as propping up specific financial markets and providing financial guarantees for the banking sector, and indirect distortions, such as government support programs for automobile manufacturers and other firms large and small. These domestic interventions, Claessens notes, also affect international capital flows and financial intermediation. Although these interventions may have been needed, governments should begin planning how to reverse the distortions. Claessens observes that international coordination is important in this respect, but extremely difficult as a practical matter.

Based on this analysis and noting that the crisis is still evolving, Claessens then draws lessons for national and international financial reform as well as lessons for emerging markets and developing countries. In regard to macroeconomic policy, he says that monetary policy should seek to address macrofinancial stability, not just price stability; governments should take advantage of boom years to reduce budget deficits; and tax systems should be adjusted so that they are no longer overly biased toward debt financing through deductibility of interest payments.

In regard to national financial architecture, Claessens specifies several key principles for strengthening regulation: it should address systemic risk and proactively identify and repair gaps in oversight and information; market discipline and supervision should complement one another; and in redesigning systems, policy makers need to be aware of the inherent limitations of financial regulation and supervision and seek to overcome them.

In regard to the international financial architecture, Claessens believes that financial information needs to be better organized and timelier. It is also important, he points out, to include non-bank financial institutions. Systems of macro-financial analysis, risk assessment, early-warning systems, and cross-border banking resolution must also be improved.

The good news of the crisis, Claessens states, is that many emerging markets and developing countries entered the crisis with more policy options at their disposal than in previous crises, since they had better fiscal health, stronger financial sectors, and better financial frameworks than ever before. Nevertheless, he says, because capital
inflows dried up abruptly at the same time as export demand collapsed, these coun-
tries still need external financing to support macroeconomic policy and social protec-
tion. In most countries, Claessens says, monetary policy should be eased and the
exchange rate adjusted to absorb the pressure. He recommends that countries in a rel-
atively strong fiscal position pursue an expansionary fiscal policy and that all coun-
tries prepare contingency plans for dealing with financial turmoil and bank failures.

Giovanni Zanalda examines two historical cases to throw light on the current cri-
sis. These two cases—the Kingdom of Naples in the seventeenth century and the
Great Depression of the twentieth—demonstrate, according to Zanalda, the impor-
tance of addressing the structure of the real economy, not just the financial factors.

In his discussion of seventeenth-century Naples, Zanalda highlights the views of
Antonio Serra, a contemporary Neapolitan physician, whose writings urged the gov-
ernment to support the real economy rather than the monetary side—the opposite of
what the kingdom was doing and continued to do in spite of Serra’s urging. In his
discussion of the Great Depression, Zanalda highlights the fateful lack of coordina-
tion among nations. In the 1930s, many countries dealt with the crisis by abandon-
ing the gold standard; however, because nations acted unilaterally without coordina-
tion, improvements in the economy of one country came at the expense of the econ-
omy in other countries.

Fortunately, says Zanalda, the current crisis is being met with broad coordina-
tion—at least so far—through the G-20, the IMF, and other means. History tells us,
according to Zanalda, that governments should consider the crisis as an opportunity
to implement structural reforms. It is also important to address the problems that
have emerged in the financial sectors in the United States and Europe. Zanalda sug-
gests that countries in East and South Asia, where much of the world’s economic
growth has occurred over the last two decades, have a special role to play in the
reform process, emphasizing the need for the government to play a greater role in the
economy, maintain stricter control on private finance, and favor the real economy
over the financial sector.

**Innovation and Competition**

Philip R. Lane examines how international financial integration affects productivity
and innovation. Lane’s focus is on innovation in developing countries and adaptation
of existing technologies developed elsewhere. Lane claims that the connection
between financial globalization and productivity holds critically important potential.
The greatest benefit of financial integration, according to Lane, is that it can raise
total factor productivity. He cites other research showing that without an increase in
productivity, the welfare benefit of international financial integration is minimal.

Lane observes that economic studies emphasize the importance of financial devel-
lopment in determining the intensity and effectiveness of innovation. But recent stud-
ies emphasize that gains from financial globalization usually depend on the level of
development and that financial globalization is not helpful if the domestic economy
is not sufficiently developed. Accordingly, Lane sets his sights on whether threshold effects are present in determining the relation between international financial integration and the level of innovation activity.

Lane says that developing countries naturally concentrate on adopting advanced technology developed abroad rather than attempting to develop new technologies themselves. Lane contends that purposeful R&D investment is required not only to develop new technology, but also to adopt existing technology. And adaptation of technology requires other things as well: human capital, institutions, and integration into the global trading system. Lane cites the work of Aghion in demonstrating that a threshold level of financial development is needed in order to finance the innovation required to absorb new technologies.

Lane reviews recent studies examining empirically the link between financial globalization and productivity, concluding that the studies—whether using firm-level, sector-level, or country-level data—show a positive relation between international financial integration and the level of productivity. Lane goes on to demonstrate, based on an econometric analysis of a collection of several data sets, that financial globalization can raise the level of innovation activity, which in turn boosts long-term productivity. Lane cautions, however, that much more research will be required to establish firmly his provisional finding.

The challenge for policy makers in developing countries, according to Lane, is to embrace financial globalization in a phased way that recognizes the interplay between domestic institutional development and greater openness to international investment flows. In terms of sequencing, the evidence that Lane provides suggests that international equity integration offers greater benefits for lower-income countries, whereas the gains from international debt integration are concentrated at higher income levels. In conclusion, Lane notes the need for internationally coordinated actions to improve the stability of the global financial system.

Sungchul Chung examines technological innovation in Korea—one of the generally recognized ingredients to the country’s astounding economic progress over the past 40 years. Chung notes that Korea’s educational preparation laid the foundation for its success as an innovator, and in the 1960s, at the beginning of the country’s rapid economic rise, educational attainment in Korea was roughly that which would be expected of a country twice as rich. Korea’s development path would not have been possible without a base of human resources capable of absorbing and improving upon the technologies transferred. Korea’s technological advance has been remarkable, but to continue to advance, it now has to strengthen basic scientific research capability and take other steps to improve innovation.

In tracing Korean innovation over the last 40 years, Chung highlights that Korean industries acquired most of their technology through informal rather than formal channels; in contradistinction to most other developing countries, foreign direct investment and licensing played only a small role in Korea. Informal channels proved much less costly and prevented Korea from falling under the dominance of multinational corporations. Chung states that the country’s industrial policy in the 1960s and 1970s essentially represented a means of learning how to absorb and improve foreign
technologies. Korea took long-term loans to finance massive importation of foreign capital goods and turnkey plants in select industries. According to Chung, Korean industries then reverse-engineered the imported capital goods in order to acquire and master new technology. In the 1970s, Korea shifted its development strategy and invested massively in the machinery and chemical industries. To help adopt the new technologies, the government created state research and development institutes, which worked with private industries to build the technological foundation for Korea’s industrial development.

Chung presents data showing that large numbers of research and development (R&D) centers sprung up in Korea in the 1980s, and Korea moved from the stage of technology learning to technology development. The government launched a national program to promote private R&D, providing tax breaks, state investments, and other incentives. By the early 1990s Korean industries had shifted their mode of technology acquisition radically from borrowing and learning from foreign sources to indigenous R&D. Investment in research and development has continued to climb, increasing more than 60 times between 1981 and 2007.

The financial crisis of 1997 hit R&D hard: R&D expenditures dropped 10 percent and R&D employment dropped 15 percent. The sector was able to recover within two years, but the financial crisis brought about two lasting changes: (1) many displaced R&D personnel established small-scale technology firms, making small and medium enterprises much more important in Korea’s R&D sector, and (2) foreign direct investment increased sharply due to the favorable investment environment created by the depreciation in local currency and asset values.

Chung states that the positive contributions of Korean R&D are undeniable, but he nevertheless enumerates several shortcomings: Korea still lags far behind the advanced industrial countries; the country does not have a well-developed university research capacity and remains weak in basic sciences; Korean R&D is highly reliant on private industry and consequently too sensitive to economic ups and downs; and Korean researchers interact only rarely with foreign scientists and institutions.
Vice President Justin Yifu Lin from the World Bank; Dr. Anne Krueger from Johns Hopkins University; Dr. Il SaKong, chairman of the G-20 Summit Coordination Committee; Dr. Euh Yoon-Dae, chairman of the Presidential Council on National Branding; Dr. Hyun Oh-Seok, president of the Korea Development Institute; and distinguished guests from home and abroad: I extend a warm welcome to all of you. This conference has been organized to look back on the development experience of East Asia and gain wisdom helpful to overcoming the current economic crisis and building a new economic framework. In just half a century the Republic of Korea achieved remarkable economic growth, named the Miracle on the Han River, transforming itself from an aid recipient to an aid donor in the global community. Korea also has a track record of successfully overcoming the financial crisis of the late 1990s and has made great strides since the outbreak of the current crisis. The signs are that Korea will be one of the first countries to emerge from it. For these reasons, it is appropriate to hold the twentieth Annual Bank Conference on Development Economics (ABCDE) in Korea.

We are being called on to overcome the economic turmoil facing us and to prepare for the new world that will emerge from the crisis. Today, I would like to talk about how we should deal with the challenges facing the global economy and what we should do to prepare for sustainable growth once the crisis has passed. The world economy is undergoing the worst recession since the Great Depression of the 1930s. Some economists say that the Great Depression was a tragedy that could have been avoided. Had there been appropriate measures and cooperative efforts at the international level such as expending fiscal stimulus and recapitalizing banks, the Great Depression might not have been so prolonged and so painful; it might have been a passing event in world history.

Jeung-Hyun Yoon is the minister of strategy and finance of the Republic of Korea.
The Great Depression taught us that it is critical to coordinate international policy and to devise appropriate economic responses if we are to defeat the economic crisis gripping the world today. In March 2008 when we learned of the potential bankruptcy of Bear Stearns, the U.S. government acted swiftly to get matters under control. Despite these efforts, financial turmoil spread to the global financial markets, which are tightly intertwined, and it soon became evident that the efforts of individual nations or only a handful of advanced countries would not be enough to overcome the crisis. I firmly believe that the two recent G-20 summits constituted turning points in the efforts to overcome the global crisis. The recent positive signs suggesting that the worst is behind us may well be attributed to the coordinated responses of the international community, such as the G-20.

Korea has been active in these international efforts. For instance, we expanded our fiscal spending in 6.9 percent of GDP and promptly lowered our base interest rate. In addition, we began to promote job sharing, a challenging initiative that is intended to encourage companies and employees to share the pain and maintain the development capacity of the Korean economy. Thanks to these efforts, the crisis affecting the Korean economy seems to be coming to an end. In fact, Korea’s GDP growth rate turned positive in the first quarter of 2009, and industrial production has been growing since January. Furthermore, the leading economic indicators have continued their upward trend.

Notwithstanding these positive signs, the private sector has not yet recovered, unemployment continues to rise, and consumption and investment have declined since the first quarter of 2008. The global economic crisis is ongoing, and uncertainties are lingering, including the potential recurrence of distress in the financial markets and prolonged recession in the real economy. These uncertainties suggest that global cooperation on macroeconomic policy is still required. Therefore, I believe that the G-20 summit should continue to play a central role in promoting international cooperation. As the chair country of the G-20 next year, Korea is committed to solidifying the cooperative framework so that the G-20 becomes the locus of international cooperation.

It is my sincere hope that this year’s conference will serve as a forum for discussing the new direction of the world economy at this historical moment of change. I expect that it will become a valuable stepping stone for seeking solutions to the economic crisis and laying out a creative blueprint for sustainable growth of the world economy. The world population, which totaled about 3 billion back in the 1960s, more than doubled to reach 6.6 billion in less than half a century. Along with the growing population, we are facing challenges that must be addressed, such as climate change, protectionism, and absolute poverty.

To address these unprecedented issues and achieve sustainable growth, a new strategy is needed. First, we must find a new engine of growth for the world economy by responding to climate change and promoting a green economy. Addressing climate change is a top priority for mankind that we can no longer postpone. It is a challenge that cannot be resolved by the efforts of one nation alone. Green growth is not a matter of choice, but an imperative; it is capable of driving sustainable growth of the
global economy based on technological advance and innovation. Last year, Korea declared low-carbon growth as our national agenda and launched efforts to turn this agenda into reality. As part of this effort, we began to promote green industries, such as new and renewable energy, and initiated the green transformation of existing industries. Korea is prepared to lead the international effort to tackle climate change and support green growth in both emerging and developing economies.

Second, we must expand international trade and investment and support world economic growth by removing all forms of protectionism. In times of crisis, governments may be tempted to raise protectionist barriers in an attempt to protect their national interests. However, once protectionist measures have been taken, they are difficult to roll back. Furthermore, they may invite a chain of retaliatory actions from trading partners that, in turn, will have a negative impact on world economic growth. As agreed at the G-20 summit, we must adhere to the grand principle of free trade. We must cooperate to remove any form of protectionism and continue our efforts to foster multilateral and bilateral trade and to liberalize investment through free trade agreements and other means. In doing so, however, advanced nations should consider and address the adverse effects of globalization, such as income inequality, on developing nations.

Third, we must work together to reduce poverty and narrow the income gap. In the wake of the current economic crisis, poverty has risen rapidly in developing countries, and the advanced nations have reduced their assistance to developing economies. Eradication of absolute poverty is a difficult challenge, but we cannot guarantee either positive progress in globalization or sustained growth of the global economy unless we address this issue. Developed countries’ support for efforts to address absolute poverty and income inequality in developing economies should be considered an investment in ensuring stable and sustained growth of the global economy. The international community—in particular, the advanced nations—should work together to address poverty by expanding development assistance to developing economies.

At this historical turning point, we are faced with the solemn task of overcoming the global economic crisis and preparing a new postcrisis economic order. Korea has a role to play in two major needs of the times. First, given our rise from absolute poverty to become an advanced developing economy, Korea intends to play a role in building a bridge between developing and advanced countries. To this end, as chair of the G-20 in 2010, Korea will endeavor to coordinate the interests of developing and advanced countries in good faith and with respect. In the Asian region in May 2009, in an effort to ensure regional financial stability, Korea helped to reach agreement on the Chiang Mai Initiative Multilateralization with China, Japan, and members of the Association of South East Asian Nations. Second, Korea will fulfill its responsibility as a global economic partner befitting its status as the world’s fourteenth largest economy. Korea will expand support to developing economies through the knowledge-sharing program, which aims to share the lessons of Korea’s development experience. Korea will also triple its official development assistance by 2015, including its contribution to international financial institutions such as the World
Bank. Furthermore, Korea will take an active approach to addressing common challenges, such as poverty and climate change, facing the world.

Winston Churchill once said, “A pessimist sees the difficulty in every opportunity; however, an optimist sees the opportunity in every difficulty.” In this spirit, we should do our best to find opportunities in the global economic crisis. One of the key objectives of this conference is to build hope for a better future. I am confident that our discussions will provide the way forward in achieving sustained growth through global cooperation.
It is a great pleasure to be in Seoul today, to address this very timely conference, and to open up our discussions on lessons from East Asia and the global financial crisis. Before sharing a few initial thoughts on the topic of the conference, let me thank our host, the government of the Republic of Korea, and our partners in this initiative for their extraordinary generosity. I also wish to thank the participants, many of whom have traveled across the globe to share their experience and ideas on how to deal with the current global crisis, which is the most serious crisis since the Great Depression.

There have been signs that the global downturn might be bottoming out: a recovery of stock markets, a decline in interest rate spreads, and improved business and consumer confidence. It is probably too early to know if some of these positive signs simply reflect the mechanic effect of the expansionary monetary policies adopted by almost all central banks and the fiscal stimulus packages under implementation around the world.

Moreover, there are also many worrisome signs on the global economic horizon. Unemployment is still rising in most economies, and capacity utilization rates are low. In most of the Western world, housing prices are still on the decline, as the oversupply of homes, tight credit conditions, and foreclosures continue to hamper the market. While mortgage rates are low and houses are, in principle, more affordable than at any time since 1980, rising unemployment is pushing more and more people to default on their home loans and foreclosures will continue to rise. The risks of a double-dip recession or a “W”-shaped recovery are still present in many parts of the world.

Even after global output growth begins to pick up, unemployment may continue to rise as capacity utilization rates remain low across the world. The International Labour Organization projects that unemployment will rise by 30 million worldwide
in 2009, of which 27 million will be in developing countries. Absent assistance, households may be forced to sell additional assets on which their livelihoods depend, withdraw their children from school, forgo necessary health care, and cut back on food, which could risk malnutrition. The effects of falling real wages and joblessness impede the ability of households to provide adequate food and necessities to their members.

Prior to the current crisis, World Bank research estimated that about 1.4 billion people were living below the poverty line worldwide. Preliminary studies of the impact of the global downturn indicate that an additional 53 million people will fall below the poverty line in 2009. The long-run consequences of the crisis may be more severe than those observed in the short run, possibly turning a short-run macroeconomic adjustment into a long-term development problem. When poor households pull their children out of school, there is a significant risk that they will not return once the crisis is over or that they will not be able to recover from the learning gaps resulting from lack of attendance. And the decline in nutritional and health status among children who suffer from reduced (or lower-quality) food consumption can be irreversible. The middle class will also be hit hard by soaring joblessness, losses in equity markets, possible currency depreciation, and anxiety over the safety of local banks.

Constrained by the erosion of fiscal space and foreign exchange reserves, many developing countries will be unable to implement countercyclical policies on their own. Moreover, the crisis is reducing their income, hereby worsening public finances, threatening existing levels of spending, and further reducing services to the poor. We may be facing nothing short of a development emergency.

In my remarks this morning, I reflect briefly on lessons learned from the East Asian experience, in terms of both economic development strategies and recovery from major macroeconomic crises. I then highlight a few issues for discussion at this conference.

Lessons from the East Asian Experience

With worsening global poverty, economists and policy makers around the world agree on at least one thing: sustainable recovery and economic development are the only solutions to the problems facing the global economy. East Asia is well placed to shed light on effective development strategies and the appropriate paths to recovery, with its long and rich history, the enormous amount of knowledge accumulated over millennia, its painful or exciting experience of booms and busts, its vibrant and diverse intellectual community, and the infinite processes of learning constantly at work.

The Recipe for Effective Development Strategies

The economic record of Asian countries over the past half-century is impressive by any measure. In 2007, the average per capita income was 3.5 times higher in South Asia than in 1960 and about 12 times higher in East Asia. In these economies in the
early 1950s, per capita gross domestic product (GDP) was less than 2,000 international Geary-Khamis dollars—as measured by 1990 purchasing power parity—and the same as in China and less than in Eastern Europe and Latin America at that time. Japan was the first success, followed by Korea; Taiwan, China; Hong Kong, China; and Singapore—the four East Asian newly industrialized economies (NIEs)—and, recently, by Thailand, Malaysia, and Indonesia. Since the 1960s, the economies of the four East Asian NIEs have maintained an annual growth rate of 10 percent for two to three decades. Such growth completely changed the poor and backward state of their economies.

Development policies that were adopted after World War II by some of these countries and led to success were inappropriate in the context of the prevailing theories at that time. Those developing countries that followed prevailing development theories in formulating their policies failed to narrow the gap between them and the industrial countries. Similarly, China’s transition to a market economy, begun 30 years ago, was thought doubtful in the light of prevailing theories. But the path taken led to sustained growth, while countries that followed standard approaches in their transitions encountered various difficulties. This contrast in economic development and transition is intriguing to economists.

According to Adam Smith and many classical and neoclassical economists, the free market is the mechanism for bringing prosperity to a nation. After World War II, and under the influence of Keynes, economists of the structuralist school challenged the free market principle. They believed that markets cannot serve as the foundation for development, because they send the wrong signals (in the presence of monopolies) or because factors of production are not fully mobile. Their ideas became the foundation of the newly established field of development economics. They believed that governments in developing countries should intervene in the allocation of resources for industrial development. This contributed to the adoption of import substitution strategies in many developing countries. The consequent debt crisis in Latin America in the 1980s and the collapse of the socialist planning system once again led to the domination of free market thinking in development economics.

The failure of the old structuralist model can be explained by the fact that many countries attempted to implement strategies that defied their comparative advantage. Their governments had to protect numerous nonviable enterprises; however, because these governments usually had limited tax collection capacities, such large-scale protection and subsidies could not be sustained with their limited fiscal resources. They had to resort to administrative measures—granting a market monopoly to the nonviable enterprises in prioritized industries, suppressing interest rates, overvaluing domestic currency, and controlling prices for raw materials—to reduce the costs of investment and operation of the nonviable enterprises. Such intervention caused widespread shortages in funds, foreign exchange, and raw materials. The government, therefore, needed to allocate resources directly to these enterprises through administrative channels, including national planning in the socialist countries and credit rationing, investment, and entry licensing in nonsocialist developing countries.
A close look at successful growth experiences of several developing countries in East Asia and other parts of the world reveals that the market was the dominant allocative mechanism in their economies; however, the state also played an active role. Moreover, the majority of industrial upgrading took place not only in exporting sectors, but also in import substitution activities (that is, import substitution that was compatible with each country’s comparative advantage). Therefore, both the structuralist and the neoclassical views may reflect important insights about the process of economic development in developing countries.

Successful Asian countries have followed economic development strategies that are consistent with their comparative advantage. In each stage of development, the market has been the best mechanism for effective resource allocation. However, economic development is a dynamic process that involves shifting from one stage to the next, which requires industrial upgrading and corresponding infrastructure improvements. Infrastructure has the character of a public good and has large externalities to firms’ transaction costs and returns to capital investment. Thus, the government should play an active, facilitating role in infrastructure improvements. The firms engaging in industrial upgrading also provide externalities to other firms in the economy. Therefore, it is desirable for the government to provide support for innovation. In developed countries, public funding for basic research and the patent system are examples of such support.

Continuous technological innovation and upgrading of industrial structures—as well as corresponding institutional changes—are the driving forces of long-term economic growth in modern times. First, the optimal structure of the economy is different for each country at different stages of development. This applies to a country’s industrial, financial, legal, and other institutional structures. Second, each stage of economic development is a point on a spectrum, not a dichotomy of economic development stages. Third, markets play a fundamental role irrespective of the stage of development. The state needs to play a facilitating role by helping to upgrade the economy from one stage of development to the next. The endowment structure, including natural resources, labor, human, and physical capital, as well as hard and soft infrastructure, is different in each stage of development and from country to country. (Hard infrastructure includes power, transport, and telecommunications systems; soft infrastructure includes the financial system and regulation, the legal framework, and social networking.)

The experience of Asia therefore suggests the need to conceptualize a new structuralist approach to development, with the government facilitating industrial upgrading by (a) providing information about new industries of comparative advantage, (b) coordinating improvements in infrastructure, (c) subsidizing activities with externalities, and (d) promoting new industries by incubation or foreign direct investment.

Like Germany, France, and other countries in Western Europe in the nineteenth century and Japan and the NIEs in East Asia after World War II, any developing country can learn from the experiences of developed countries in technology and institutions. They can undertake rapid technological improvements, upgrade their industry, and adapt institutions at a relatively low cost and with less risk. Such a
strategy could allow them to maintain rapid economic growth for several decades, narrow the gap with developed countries, and even overtake some of them.

**How to Recover from Economic Crises**

There are some similarities and some differences between the 1997–98 East Asian crisis and the current global crisis. In terms of similarities, both crises were made possible because the volume and patterns of international financial flows have increased considerably in recent decades. This evolution was underpinned by many factors: the deregulation of markets, which lifted capital controls in many developing countries, de facto or de jure; the high returns to portfolio investment in East Asian financial markets; and the improvement in the general economic outlook and the new developments in the technology of international financial transactions.

While East Asian countries benefited enormously from their integration into the global economy, they also realized that globalization and liberalization carry important risks. In Thailand, where the 1997–98 crisis was ignited, investors, who had been encouraged by high growth rates and had overlooked weaknesses in the financial and corporate sectors, suddenly lost confidence. Policy makers in these countries had difficulty addressing the problems of an overheated economy and the weakening of the current account. This created doubts in international financial markets about the compatibility of the monetary and fiscal stance with an exchange rate tightly linking the domestic currency to the U.S. dollar. Such inconsistencies in macroeconomic policies were a clear recipe for capital flight.

Another similarity with the current global situation is the fact that the East Asian crisis revealed inadequacies in the management, supervision, and regulation of financial institutions. Flooded with liquidity, these institutions had accumulated large amounts of risky assets and contingent liabilities against which they had inadequate capital and reserves. Thus, the crisis revealed the need to update regulation, improve transparency and supervision of financial institutions, and rethink existing frameworks for dealing with troubled banks. It also raised concerns about the effectiveness of rating agencies and deposit insurance schemes. It now appears that these emerging-country problems were not paid sufficient attention in industrial countries.

Just like the Asian crisis, the current downturn has had repercussions across countries, with some economies suffering from severe capital outflows and declining stock market and asset prices because of contagion. As was the case about a decade ago, economic contraction is reflected in the decline in exports, which heightens risks of the loss of creditworthiness.

But the current global crisis, which Alan Greenspan recently called a “once-in-a-century credit tsunami,” differs from the East Asian crisis in several respects. First, it originated in the U.S. and European financial systems. The bursting of the U.S. tech-stock bubble in 2000–01, which had a substantial effect on the wealth of American households, forced the U.S. Federal Reserve to aggressively ease monetary policy. It lowered either the federal funds rate or the discount rate 27 times between January 2001 and June 2003, with the funds rate falling from 6.5 to 1.0 percent over that
period. This expansionary monetary policy averted a deeper recession by stimulating a boom in the housing market, which soon turned into a housing bubble. Higher housing prices fueled a consumption boom, and the Fed’s continued expansionary monetary policy kept the U.S. economy awash in excess liquidity. Another essential ingredient behind the persistently low U.S. (and global) real interest rates was the shift of developing countries toward accumulating large volumes of U.S. assets, motivated by their experience in previous crises and made possible by their current account surpluses. This made it possible for the United States to finance its massive current account deficit over a prolonged period without abrupt changes in real interest rates or real exchange rates.

Second, the high levels of financial innovation in major financial markets, driven by a search for higher yields in a low-interest-rate environment, compounded the levels of indebtedness in risky assets of major financial and nonfinancial institutions. Much of this innovation was carried out through excessive leveraging by firms whose activities were not regulated and other new instruments that were too complex to be regulated effectively. As a result, policies tended to advocate for deregulation of financial markets and were sometimes accompanied by lax supervision.

Third, contrary to the Asian crisis, policy responses were bold, comprehensive, and countercyclical. Their goal was not only to strengthen aggregate demand but also to maintain the availability of credit to households and businesses. This was in sharp contrast to what occurred in Asia in 1998, where the high domestic interest rate policies adopted to encourage the retention of resources in national economies initially attracted further capital inflows and external borrowing by domestic residents. As a consequence, financial institutions and the private sector assumed growing levels of foreign currency risk, which eventually made East Asian economies more vulnerable to external shocks. This was not the case this time. Fiscal policy and monetary policy, as well as government guarantees and safety nets, such as deposit insurance, have been used quickly in almost all industrial countries to improve conditions in the financial sector. Moreover, most East Asian economies have had strong fundamentals and vibrant exports in the years preceding the current crisis.

East Asian countries were able to recover quickly from the crisis because of their generally good macroeconomic fundamentals, the result of successful development strategies implemented for several decades; the good global environment prevailing at the time, which allowed many countries to regain their growth momentum through exports; and the rapid implementation of broad-ranging financial and corporate sector reforms such as the recapitalization of banks and closure of insolvent financial institutions, with their assets transferred to a restructuring agency (Korea, Indonesia, Thailand); the strengthening of prudential regulations, including loan classification and provisioning requirements and capital adequacy standards; the introduction of more stringent conditions for official liquidity support (Indonesia, Malaysia, Thailand); the adoption of competition and governance policies and trade reforms (Indonesia, Korea); the implementation of social sector policies such as labor-intensive public works programs (Indonesia, Thailand) and expansion of the unemployment insurance system (Korea); and the provision of higher public spend-
ing for health and education (Indonesia) and the reallocation of budgetary expenditure to health programs for the poor (Thailand).

Many of the broad lessons learned from the Asian crisis are still valid: first, the importance of the credibility of macroeconomic policies, both to domestic agents and to external lenders; second, the need to ensure that financial sector liberalization goes hand in hand with the adoption of a strong framework for prudential regulation and supervision of financial institutions; and third, the necessity to monitor and manage carefully capital inflows and their potential disturbances.

However, East Asia’s most valuable lessons for dealing with today’s crisis may be the willingness to consider carefully designed, unorthodox policies, especially in difficult circumstances. These lessons are illustrated by the contrasting paths chosen by Japan and China to fight deflation.

Japan’s story of the 1990s, often referred to in the economic literature as the “lost decade,” is well known. After the financial deregulation in the late 1970s, Japanese corporations were allowed to raise capital more cheaply from the capital markets, especially the foreign bond market, and savers could invest in the equity markets. In the second half of the 1980s, Japan’s monetary authorities flooded the market with liquidity in order to enable businesses to cope with the rising value of the yen. A high savings rate and strong export surplus contributed to the ease of credit. Large excess capacity built up during the decade.

Fierce competition and decline in margins forced banks to increase their risk profile. Businesses invested in new capital equipment in an effort to become more competitive in international markets, but the excess liquidity also found its way into speculation in Japan’s stock market, in real estate ventures, and in foreign investments. Between 1980 and 1996, loans to the real estate and construction sector nearly doubled from 11 to 19 percent, loans to the financial sector tripled from 3 to 10 percent, and loans to the manufacturing sector dropped from 32 to 15 percent.

The disproportionate growth of loans to real estate and the financial sector resulted in a large run-up in equity and real estate prices, which led to a buildup of asset bubbles. With the burst of the asset bubble, the Nikkei declined by one-third within a year from the peak at the end of 1989 to the end of 1990. Banks suffered huge losses from nonperforming loans. The collapse of large financial institutions in the late 1990s increased the cost of finance, and the credit crunch further caused a contraction in the economy. During the “lost decade,” Japan’s government implemented very aggressive fiscal stimulus policies. In 1991, public debt represented 60 percent of the country’s GDP. By 2002, it had increased to about 140 percent—implying a very large and very decisive stimulus of 7 percent of GDP per year. Yet Japan did not get out of the crisis. Households responded by saving more, which dampened the effects of government spending.

By contrast, China chose a completely different path when its economy entered a five-year period of deflation at the end of 1997. In the midst of the Asian financial crisis, neighbor countries all depreciated their currencies, with sharp economic slumps in Indonesia, Korea, Malaysia, the Philippines, and Thailand. There was heavy deflationary pressure on China, with many economists advocating a parallel depreciation of the renminbi. Instead, the authorities decided to hold the exchange
rate steady and to orchestrate a very large (about $1 trillion) fiscal expansion in 1998–2002. The government issued an estimated ¥ 660 billion in bonds specifically to finance infrastructure, which may have induced four times more bank loans and private and local government investment. Thanks to decisive government efforts to increase public investment, the ratio of total investment to GDP remained high at 33 percent in 1998–2002. As a result, the network for land transportation increased from 1.08 million kilometers in 1990 to 1.47 million kilometers in 2000, before jumping to 3.53 million in 2006. Following the same trend, the total transportation network initially increased by two-thirds, from 1.55 million kilometers in 1990 to 2.61 million kilometers in 2000, before almost doubling to 4.85 million kilometers in 2006. That strategy allowed China to combat deflation, with uninterrupted growth at 8 percent on average during that period.

The Chinese economy emerged from deflation in 2003, and the average annual GDP growth rate reached 10.8 percent in 2003–08. Due to the higher growth rate, the government’s revenue grew rapidly, and the public debt declined from about 30 percent of GDP in the 1990s to about 20 percent in 2007. The improvement in infrastructure facilities, especially in rural areas, provided a sounder basis for economic growth. The fiscal resources spent on infrastructure investment led to higher growth that was sustainable.

The success of China’s economic stimulus shows not only that well-designed stimulus can enhance growth in the developing world, but also that there is room for this growth-enhancing bottleneck-releasing type of investment. In contrast to Japan, where ineffective government stimulus investment and limited investment opportunities led to a sharp increase in government debt in its lost decade, China’s effective stimulus led to higher growth, higher fiscal revenue, and lower debt.

This suggests that, if the resources are used appropriately, loans can be repaid from the high return on these investments. It is therefore important for the developing world to have access to the resources needed to join a global effort at shortening this crisis. Transferring resources to the resource-constrained developing world to invest in bottleneck-releasing and growth-enhancing efforts is a positive sum game, enabling developed countries to seize the high-return investment opportunities and, at the same time, allowing developing countries to realize higher growth in the long run, while stimulating demand in the short run.

With the existence of excess capacity in the global economy, solving the problems at hand would require much more than traditional Keynesian economics. While the restructuring of credit and financial markets in developed countries was a necessary first step that also helped to avoid a currency crisis in developing countries, efforts to increase bank lending may be futile because of firms’ lack of good investment opportunities and households’ lack of confidence about their future job security. The consequent increase in nonperforming loans and the risk of more assets becoming toxic could jeopardize efforts to stabilize the financial sector. Even if it were possible to restore confidence within the financial markets and to unclog the channels of credit, increasing money supply by keeping interest rates low might not stimulate demand.
Excess capacity in developed economies implies limited profitable investment opportunities, pessimistic expectations, low confidence about the future, and the likelihood of a liquidity trap. When confronted with excess capacity, companies cut prices to reduce their excess inventories. This aggravates the slack in labor markets, which in turn leads to decline in wage growth and prices. Even when nominal interest rates are low, as is currently the case around the world, investment opportunities are limited by excess capacity. The ensuing fall in demand and prices is self-reinforcing: the increase in unemployment and decline in wages result in a further fall in demand and prices, and the increase in the real value of nominal debts can create severe problems of default, which may exacerbate credit losses of financial institutions and further result in a vicious cycle of debt deflation.

Dealing with this large synchronized financial crisis is beyond the capability of any single country. To overcome this synchronized financial crisis, decisive and concerted efforts are needed. High-return public infrastructure investment opportunities may be limited in mature economies. By contrast, they tend to abound in developing countries, as infrastructure is the main bottleneck to growth. Release of evident bottlenecks in developing countries would result in enhanced growth potential, higher marginal returns to private sector investment, and higher government revenues to pay for the projects. While such bottlenecks are the best investment target for effective fiscal stimulus, many developing countries are constrained by their fiscal space and availability of foreign reserves. These constraints bring into question the feasibility of the traditional Keynesian macroeconomic policies.

Loans from developed to developing countries to finance high (economic and financial) return projects can lead to a win-win situation, as they will enhance the growth potential and increase sustainable demand in the long run. By supporting investment in infrastructure bottlenecks that constrain economic growth in developing countries, the coordinated fiscal stimulus would have a large multiplier effect, raising the chances for the global economy to avoid remaining trapped in a downward deflationary spiral.

A Few Issues for Discussion

I now turn to the subject of this year’s conference: lessons from East Asia and the global financial crisis. Beyond the painful cruelty of the crisis (loss of wealth, unemployment, and the social and political consequences), economists should welcome the fact that the crisis is likely to spark an evolution in thinking about economic development—its nature, its causes, and the choice of policies it requires. Despite the unpleasant criticism and the sometimes unfair challenges to our discipline, economists must acknowledge some mistakes—such as the naïve belief in the end of volatility and complacency—and gaps in the existing stock of knowledge. We must use the current crisis as an opportunity for new thinking on macroeconomic and development issues. Specifically, there is a need to undertake the following:

• Revisit some of the dominant analytical frameworks in macro and development economics
• Draw lessons from the past crises and the various experiences and policy responses of East Asian countries

• Undertake new theoretical and empirical investigations that may be needed to expand the stock of knowledge

• Explore alternative methodological approaches that would complement (or substitute for some of) the existing ones.

We have an exciting and wide agenda ahead of us, and a vast range of issues deserves our attention. Let me flag a few possible topics that are likely to generate a lot of discussion over the next couple of days.

The first is the reorganization of financial and capital markets. The global crisis has shaken the confidence in financial markets of investors and policy makers around the world. It has also sparked a discussion on how best to reorganize and regulate markets so that they can effectively mobilize and allocate savings among competing uses and help firms to create value. In an effort to contain the crisis, the authorities in the United States and many European governments have taken the unprecedented steps of providing extensive liquidity, giving assurances to bank depositors and creditors that include blanket guarantees, structuring bailout programs that include taking large ownership stakes in financial institutions, and establishing programs for direct provision of credit to nonfinancial institutions. Some developing countries have reintroduced capital controls to prevent capital outflows. Many old and new questions are back on the agenda:

• What is the proper role of government in the financial sector?

• What is the optimal financial structure for developing countries? Should there be different models of financial development for countries at different levels of economic development?

• What is the proper policy on capital account liberalization, and should there be different policies for different types of financial inflows (portfolio investments versus foreign direct investment)?

• When and why should government bail out or own financial institutions? Are blanket guarantees the most efficient instruments to halt a systemic crisis?

• How should prudential regulation be designed and implemented to facilitate monitoring, supervision, and innovation?

• What is the proper policy on antitrust and bankruptcy legislation?

Second is the role of government and first-order principles of good economic policy. The global crisis has not changed the basic objectives of economic policy: prosperity, equity, and stability or continuity. But it has challenged the general premise that markets tend to deliver socially superior outcomes. The boundaries and balance between markets and the state and the central role of institutions (regulations and regulatory frameworks in the case of financial markets) will be at the center of the debate in development economics in years to come. Even in the context of industrial
countries, the most successful economies are not those that minimize the role of government, but those that adopt a pragmatic approach to economic policy and have the government intervene when this intervention produces superior social outcomes—even if doing so implies higher levels of taxation or management by the state of pensions, medical insurance, and so forth. This observation makes the following questions even more relevant:

- Under what conditions are the traditional broad principles of good macroeconomic management (free markets and sound money) conducive to sustained growth and poverty reduction?
- Even if these broad principles still apply to all countries in an increasingly globalized world, how should they be operationalized, especially from the perspective of developing countries?
- What new principles of economic policy can be learned from the experience of East Asian and Latin American countries that seem to have weathered the storm?
- Given that sustained growth is mostly about continuous industrial and technological upgrading, how should industrial policies be designed and implemented?
- How should the comparative advantage theory (Lin 2009) be integrated into development economics?
- What institutions really matter for growth, especially from the perspective of developing countries? What have we learned about the virtuous circle between institutions, social capital, and economic development?

Clearly, we will not be able to settle these questions—and many others on the agenda—over the next two days. Still, given the impressive collection of great minds attending this meeting, I am confident that we will make serious progress and, perhaps, enrich development thinking.

In closing, let me reiterate the World Bank’s commitment to cutting-edge research that is also highly relevant to development policy. We are currently examining the future directions of our research and will welcome findings and recommendations from this conference. We will also work to improve outreach, dissemination, and access and to strengthen collaboration with local partners around the world.

References

Good morning. I am thankful to the World Bank and the Ministry of Strategy and Finance for giving me the opportunity to speak before this distinguished audience.

Although there are signs of green shoots, the global economy is still in the midst of the worst crisis since the Great Depression of the 1930s. Some scholars even describe the current global financial and economic crisis as “depression-sized” because of its severity and magnitude.

This morning I would like to draw your attention to the major causes of the crisis and how the global community as a whole has been responding to it. I am of the view that if there were no such globally concerted efforts, the crisis could have become another depression, not just a depression-sized recession.

According to the National Bureau of Economic Research (NBER), the average duration of postwar recessions in the United States was 10 months, and the two longest recessions lasted 16 months each. Both of these recessions occurred after the two oil crises in the 1970s and the 1980s. The NBER officially declared that the current recession started in December 2007, so it has already broken the previous record for duration.

However, with the globally concerted policy efforts that have been undertaken, I am sure that a global economic recovery, although anemic, will begin soon, perhaps in a few months. My discussion of the global policy responses is based primarily on the G-20, which produces more than 85 percent of global gross domestic product.

Let me now turn to the causes of the current crisis. Indeed, many factors are blamed. Some observers point to the lack of prudent regulation and supervision. These critics go so far as to say that the current crisis was caused by “the greatest regulatory failure of modern history.” Still others point to excessive greed, excessive risk
taking, and overleveraging on the part of financiers. Others single out low interest rates, overly accommodating monetary policy, and global imbalances. Certainly, no one factor deserves all the blame. These factors are all interrelated, and they are both the causes and the consequences of each other. Therefore, it is difficult to single out the main causes of the crisis. I am sure that numerous dissertations, articles, and books will be written on the subject in the years to come.

That said, let me try to explain how the crisis unfolded, starting from low interest rates and accommodating monetary policy with excessive liquidity. Low interest rates together with excessive liquidity pushed the financial markets to search for high yields. Since financial yields are inversely related to the quality of credit, the result was a boom not only in subprime mortgages but also in other risky assets. Particularly in the absence of a prudent regulatory and supervisory regime for the whole financial sector, shadow banking and off-balance-sheet transactions multiplied. Also, with rapid digitalization and computerization, new financial instruments were introduced so quickly that the regulators and supervisors could not keep up with developments.

Obviously, these developments could not be sustained without the existence of global imbalances of savings and consumption or the “global savings glut,” which supported the low interest rate regime in the United States for so long. Furthermore, with accelerated globalization—financial globalization in particular—the crisis started in the United States and immediately became a global crisis.

Naturally, the overleveraging, excessive risk taking, and excessive spending of U.S. consumers were doomed. When the subprime mortgage market started to crumble, the negative impact spread quickly throughout U.S. financial markets. You know the rest of the story: the contagion spread throughout the world, making the crisis the worst since the Great Depression of the 1930s.

At the early stage of the current crisis—that is, when the U.S. government dealt with the Bear Stearns problem in March 2008—none of us foresaw the severe global crisis in the making. Many saw the problem as specific to one nation or one sector. Consequently, policy responses at the time were based mostly on initiatives in individual nations.

However, following the bankruptcy of Lehman Brothers in September 2008, the world had to come to terms with the seriousness of the crisis and to recognize that combating it required concerted global responses. Leaders of the 20 globally important economies gathered together in Washington, DC, on November 15, 2008. In my view, the gathering itself was a historic event for global governance. Until then, under such circumstances, only leaders of the G-8, which has been acting as if it were an informal global steering committee, would have gathered. This time around, given the global nature of the problem, leaders of the systemically important emerging economies were also invited to the Washington meeting.

On the day of the Washington G-20 summit, the Washington Post described the occasion as marking a “historic power shift.” To be more accurate, I would say that it was a delayed recognition on the part of the G-8 of the shift in global economic power that had been taking place during the last couple of decades. For the same reason, Paul Krugman claims, “This is two decades too late. But better late than never.”
In any case, the G-20 leaders in Washington, DC, shared a common view regarding the seriousness of the current crisis and agreed to take concerted policy measures. This global response would differentiate the current crisis from the Great Depression of the 1930s. You will remember that the world leaders met in London in 1933 with a purpose similar to that of the gathering of G-20 leaders. However, the meeting in 1933 failed to agree on exchange rate policies. In stark contrast to that, the G-20 leaders in Washington last November decided to meet again on April 2, 2009, to follow up on their agreements and to see if additional measures were needed. As you all know, the third G-20 summit is scheduled to be held in Pittsburgh on September 24–25, 2009.

Now, I would like to bring your attention to the G-20 summit meetings held in Washington, DC, and London in November 2008 and April 2009, respectively, and outline briefly what the leaders have achieved so far.

First of all, the leaders in Washington, DC, formed a consensus to have broader concerted policy responses. They agreed to use fiscal measures to stimulate aggregate demand rapidly, although they did not issue specific guidelines. Second, they also decided to resist protectionism and to avoid repeating the same mistake the world made in the 1930s. More specifically, they agreed to have a “standstill” agreement on protectionist measures regarding trade and investment for at least one year.

At this juncture, critics may point to the fact that by March 2009, 17 out of 20 countries had taken 47 protectionist measures in one form or another. They would say, “What is the big deal about a standstill agreement that does not do any good?” My response would be that none of those measures has seriously distorted global trade yet. It is my view that given their public commitments, the leaders would pause before taking any serious protectionist measure.

The leaders in London this year committed to monitoring and publicly reporting protectionist measures for the purpose of “naming and shaming” effects. In addition, the leaders in London recommitted to the standstill agreement and even the rollback of the protectionist measures taken since the Washington meeting. By the way, it was President Lee Myung-bak who successfully advocated both the standstill and the rollback ideas.

Another critical achievement of the two G-20 summits is that special attention was paid to emerging and developing economies, which suffer the most in times of crisis. The leaders’ decision in London to enhance international financial institutions’ resources with an additional $1.1 trillion would benefit primarily developing and emerging economies. The rest of the world would also benefit, as these resources would become an additional global stimulus as they are spent.

Let me now take a brief moment to say a few words on the role the Republic of Korea played in the process of preparing the London summit. First, Korea made strenuous efforts in cooperation with the United Kingdom, the chair country for the London summit, to have the G-20 leaders agree to produce deliverables and implementable measures. Toward this end, Korea played a key role in bringing the International Monetary Fund (IMF) and the World Trade Organization (WTO) into the G-20 summit to take advantage of their expertise and analytical support.
Second, Korea tried its best to bridge the gap between the industrial countries and the emerging and developing economies and to support the causes of these economies. Undoubtedly, we Koreans still have a vivid memory of the pains of absolute poverty and the process of development. Also, Korea, which succeeded as a country primarily through international trade, knows the value of free trade better than anyone. This is well reflected in President Lee Myung-bak’s trade policy initiatives at the two summits.

Now, I would like to turn to the prospect of the forthcoming G-20 summit. With the globally concerted policy measures taken by major global financial and economic players, the recovery of the global economy appears imminent. However, considering the fragility of the recovery, it would be premature to put the “exit strategies” outright on the agenda for the Pittsburgh G-20 summit in September 2009. Doing so would most likely give the wrong signal to the market, pouring cold water on the green shoots of recovery. Therefore, I would rather see the exit strategies fully discussed at the fourth G-20 summit, assuming that the leaders will decide to meet again sometime next spring. In fact, for this very reason, the world may need another G-20 summit meeting after the Pittsburgh summit.

At this point, I would like to reemphasize that the G-20 leaders should be mindful of the dire consequences of premature exit policies taken during the 1930s in the United States and the 1990s in Japan. Considering the “depression-sized” current crisis, the G-20 leaders might want to take the risk of unwinding their countercyclical policies later rather than sooner. I agree with Martin Wolf, who says, “The world needs aggressive monetary and fiscal policies far longer than many believe.” Of course, there is the danger of “malign stagflation” if the global stimulus is withdrawn too late. However, considering the pace at which private sector balance sheets are being repaired all over the world, the likelihood of stagflation does not seem to be higher than the danger of reverting to recession if exit strategies are implemented too early.

Clearly, at the Pittsburgh summit on September 24–25, 2009, the leaders will have to follow up on their London agreements by making sure that they are being implemented properly. You will remember the cheer the global community—the financial community in particular—gave to the outcome of the G-20 summit in London. I believe that this was because the summit came up with not just rhetoric but also deliverables. However, to make the G-20 summit effective and relevant, the deliverables need to be delivered in due course. I would like to emphasize again that in the process of following up, special attention must be given to developing and emerging economies.

The leaders also spent a substantial portion of their time at both summits in discussing the reform of the international financial system to prevent the recurrence of similar crises in the future. In doing so, the G-20 leaders, instead of overhauling the existing international financial architecture, decided to improve on the existing international financial institutions. In other words, instead of aiming for a Bretton Woods II by creating new supranational financial institutions, they agreed to strengthen the Bretton Woods institutions and the Financial Stability Forum. To enhance the political legitimacy and operational effectiveness of these institutions, the G-20 leaders were particularly mindful of emerging economies’ increased representation and enhanced voices.
In my view, their approach is not ideal, but it is realistic and practical, considering the complex global decision-making process. However, at the forthcoming and future G-20 summits, bolder reforms regarding the governance and mandates of these institutions should be introduced to make them more legitimate and effective.

For example, with regard to governance of the IMF, in addition to the agreement on the expedited quota adjustment made in London, representation on the executive board needs to be adjusted further. Currently, eight out of 24 board seats are held by European countries, which is disproportionately larger than their current economic weights. To make the IMF more reflective of the global economic reality, readjustment seems to be the order of the day. Doing so would make the IMF more legitimate and effective.

Also, as was agreed in London, the enhanced IMF surveillance in cooperation with the newly established Financial Stability Board (FSB) should be applied even-handedly to both industrial countries and the rest of the IMF members.

At the same time, it is important to increase the role of the IMF in overseeing the international capital markets and their transactions. These issues should be addressed fully in the Pittsburgh summit.

Although the global imbalance is considered to be one of the major fundamental causes of the current crisis, this issue was not fully debated in the two previous summits. However, at the Pittsburgh summit the leaders should earnestly tackle the issue for its early resolution.

In addition, as unemployment may steadily increase for some time to come, the G-20 leaders at the Pittsburgh summit should recommit themselves to resisting protectionist pressures.

Even with these efforts at both the global and national levels, there still may be a gap in the provision of public goods for sustained global financial stability and economic prosperity. Taking this into account, the global community should encourage rather than discourage regional cooperation to fill the gap. In this light, the recent expansion of the Chiang Mai Initiative and other regional arrangements should be welcomed, provided their mode of operation remains consistent with that of the existing multilateral organizations.

Another critical issue that should be put on the agenda of the Pittsburgh summit, in my view, is how to strengthen global governance by making the G-8 and the G-20 complement one another. Although some say, “The G-8 is obsolete,” and others say, “The G-20 will sound the death knell of the G-8,” I do not think that the G-20 summit should intend to replace the G-8 summit. Instead, there should be a serious political endeavor by the G-20 leaders to find ways to cooperate and coordinate for maximum synergy between the two processes. I personally think that with an appropriate division of work and mechanism of cooperation, they can be mutually reinforcing of better global governance.

The real challenge is how to work out the appropriate division of work and cooperation. No doubt this will be a politically complex and sometimes painful process. Even then, the G-8 leaders together with the rest of the G-20 members should make conscientious efforts to attain this goal. I was glad to see that the recent meetings of G-8 finance ministers in Italy discussed cooperation with the G-20. I hope that the meeting of the G-8 leaders in L’Aquila in July will keep the momentum going.
Before concluding, let me say that the current economic and financial crisis proves yet again that the world is ever more interdependent and more integrated; the problems not only of neighbors but also of a far distant country can quickly become one’s own and those of the world. In light of this, in addition to appropriate policy measures and institutional reforms at the national level, the G-8 and the G-20 together on behalf of the whole global community should be responsible for appropriately managing and strengthening global governance with the aim of achieving sustained global stability and prosperity.

Once again, I thank you for inviting me to be part of this conference.
The world economy has been experiencing its worst shock since the Second World War. In both depth and breadth, the world has not experienced anything even close to this severity. All countries have been affected except those entirely cut off from the international financial system, and those are the countries whose per capita incomes are low and stagnant.

Unlike earlier crises, industrial countries are fully part of it. Developing countries are experiencing a downturn not of their making (but that was also true of earlier recessions). There are a number of questions about how developing countries may need to adapt their policies to foster economic growth in light of the crisis. Various prognosticators have argued that the financial crisis shows that outward-oriented development strategies are a thing of the past, that the crisis proves that market-based economic policies are misguided, and that new development strategies are needed.

But these conclusions are, in my judgment, in error. In the future, more attention will certainly need to be paid to the financial system and an appropriate regulatory framework, capital will likely become more expensive, and these factors alone will reduce the latitude for misguided policies in the future. However, that is not a prescription for changing economic policy more generally.

To develop the argument, several issues need to be addressed: (l) the role of the financial system, (2) the outlook for the international economy in the near, intermediate, and long term, (3) the way in which different countries and groups of countries have been affected, and (4) on the basis of considerations for 1 and 2 and 3, the lessons for the future.
The Role of the Financial System

In assessing policy options and prospects, a first point to remember is that a well-functioning financial system is essential for economic growth and that the more advanced the economy, the more important it is to have a deep and well-functioning system. The vital economic function played by the financial system is to evaluate the risk-reward trade-offs of alternative investment opportunities, selecting those that have the highest risk-adjusted payoffs and enabling different groups of savers to allocate their resources to those investments closest to their own risk-reward preferences.

Any economic history of the industrial countries must include a record of the development of the financial system alongside the real economy. Starting from a world in which virtually all investments were owner or family financed (with a consequent failure of highly profitable firms to grow as rapidly as they might and less profitable firms to grow more rapidly than they should), short-term finance developed in a variety of ways, among them sharing among guild members and early bank financing of commercial paper. Longer-term lending, equity finance (which enabled the development of large-scale enterprises, today’s modern corporations), start-up investors (venture capitalists and angels), and so on all became part of the financial intermediation industry. None of these could have happened without the development of a commercial code and an appropriate set of institutions to enforce contracts and settle disputes.

Numerous studies have documented the strong relationship between level of per capita income and extent of development of the financial system. In considering how national or international financial architecture might be altered, it is important to remember that the risks in the financial system can often be diminished, but they cannot be negated entirely without negative consequences for economic growth. Regulation is needed, but the wrong sorts of regulation or overregulation may lead to stagnation or relatively slow growth. The art is to find the happy medium.

However, it has long been recognized that financial intermediaries are “special” in various ways. An important one of those ways is that many financial institutions, and especially banks, are lending long and borrowing short. There is an inherent mismatch in maturities between their assets and their liabilities. Because the future is inherently uncertain, lenders will not lend without receiving an adequate reward for the risk they are taking. This gives rise to the “risk premium,” which borrowers pay to compensate lenders for accepting the risk.

This can lead to perverse incentives in difficult situations. If a bank recognizes, for example, that its equity is or will be greatly eroded, there is a strong incentive to “gamble for resurrection” by making highly risky loans. Such loans naturally bear higher interest rates because of the risk. If the borrowers are able to repay, the bank is “saved.” If the borrowers are unable to repay, the bank owners are no worse off, because the bank would have failed anyway.
Thus, there never was a strong case for totally free financial intermediation, unless incentives could be found that would induce bank owners and managers to avoid taking on too much risk, especially if they encountered adverse fortunes. The requirement that banks (and other financial institutions) maintain adequate equity relative to the riskiness of their loans is set precisely in order to avoid the perverse behavior that might otherwise arise if banks face difficulties.

However, regulation itself is difficult and imperfect. The risk classifications in Basel I turned out to be ill-advised in several ways, including the categorization of various forms of loans into risk categories. And Basel II seems to be undergoing great scrutiny. The art of regulation is to thwart the perverse incentives that might arise, but simultaneously to permit the development of an effective and efficient financial system for evaluating risk-return trade-offs and appropriately allocating loanable funds. Since a considerable fraction of loanable funds is devoted to increasing the nation’s capital stock, the efficiency of the financial system is an important determinant of the rate of economic growth.

That said, there will always be financial crises. Perhaps the first lesson of this (as well as past) financial crisis is that it would be foolhardy to expect that it would never happen again. Even so, the timing, proximate causes, and nature of any future crisis will be different from those of the past. The best that individual countries can do is to develop their own financial system in accordance with good regulatory practice, while supporting the development of international norms, regulations, or cooperative regimes to address some of the international aspects of financial regulation.

While financial crises are painful, and we must seek to understand their causes and implement policies that will reduce, if not eliminate, their negative impact, it should also be remembered that there are worse things than crises. Researchers have shown that countries that have had financial crises have, on average, grown more rapidly than countries that have not; the reason is that the financial systems in the crisis countries, although imperfect, generated more credit and a higher rate of return than those in countries where there was financial stability through heavy regulation, with resulting lower growth rates (Rancière, Tornell, and Westermann 2008).

The point is dramatically illustrated by the comparative experience of the Republic of Korea and India. Figure 1 shows the income per capita of Korea and India in the period since 1960. The two countries started at about the same level of income per capita, but economic growth was considerably faster in Korea than in India. Korea, however, had the major crisis of 1997, which shows clearly on the figure. India did not have such a crisis, as shown by the continuous (but slower at least until the 1990s) upward slope of the line for Indian income per capita. Forced to choose between the Korean growth rate cum crisis and the Indian growth rate sans crisis for the next half century, most people would choose the Korean experience for their countries. Thus, while measures to reduce the magnitude of crisis are urgently needed, care must be taken so that those measures do not so stifle the financial system and greatly reduce growth rates.
In this regard, one additional lesson, which is valid for all countries, is worth noting. Until this crisis, the tendency among policy makers and academics had been to regard finance as a field separate from macroeconomics. One of the fundamental lessons of the crisis is that finance is an integral part of macroeconomics and vice versa (how else could the real economy have been so affected by the financial system?). It seems virtually certain that over the next few years, many economists and finance specialists will be trying to understand better the interrelations between finance and macroeconomics.

**Causes of the Crisis**

A first question pertains to the causes of the crisis. There will surely be analyses and studies for a long time to come, as researchers seek a better understanding of what happened and why. At present, there is broad understanding of the following features.

First of all, there are several levels of analysis. While there is little doubt that the crisis was financial, it was, and is, “real.” The first apparent difficulties took place in the housing sector, and concerns about the degree to which home loans would become nonperforming in various portfolios marked the beginning of the crisis (concerns really covered the actual rate of nonperforming loans and the uncertainty about who held how many). But the factors that immediately led to crisis—the housing bubble (in some countries), then the mortgage crisis, then the financial crash—themselves had causes.
That immediately raises the question, Why was there a housing bubble? One factor, certainly, was low real interest rates. The years 2002–07 saw the most rapid period of inflation-free economic growth in human history (except, perhaps, in postwar reconstruction or recovery from natural disaster). One of the reasons for that growth was that the world had experienced very low real interest rates. The demand for housing ownership is highly sensitive to the real interest rate, and low real interest rates were a major factor in the housing boom.

In addition, low real interest rates led to a “search for yield,” as investors sought a higher return than was available with safe assets. That led many portfolio managers and individual investors to increase the share of their assets invested in emerging markets (where spreads were inordinately low) and other risky ventures, which also contributed to the financial crisis.

The fact that low real interest rates were a major factor encouraging the housing bubble and investment in risky assets more generally raises the next question, Why were real interest rates so low? At that deeper level, much of the blame must, in my judgment, go to global imbalances. As everyone knows, the United States ran large current account deficits during the years prior to the crisis, while China had large current account surpluses (to be sure, oil exporters and a few others contributed significantly to the surpluses, but China and the United States were the major players). While focus has been on the American deficits, it seems clear that if the United States had not run large current account deficits, either there would have been a global recession earlier or some other countries would have had to make their economic policies more expansionary. In the longer term, unless the world adopts a system, probably rules based, for the resolution of unsustainable global imbalances, the same sort of difficulties will arise again at some (reasonably distant) future date. But for the next decade, at least, it is unlikely that a buildup will occur again (although there may be other problems that I will discuss).

**Outlook for the International Economy**

It is obvious that the prospects for developing countries and emerging markets will be better the healthier is the growth of the international economy. Here, there are two questions. First, what sort of economic growth will emerge over the medium and longer term? Second, and not totally unrelated, to what extent will the trend toward increasing openness and deeper integration of national economies continue?

Turning to the first question, many observers have questioned whether the industrial economies can resume the rates of economic growth that they enjoyed in the two decades prior to the crisis. There are several reasons for skepticism: (1) concern that consumers and firms will want to restore their balance sheets by saving much of any increase in income when recovery does start; (2) belief that part of earlier growth was unsustainable and cannot be repeated; and (3) belief that the focus on fiscal deficits and buildup of government debt in the industrial countries may “crowd out” private investment and therefore impair growth prospects.
The first concern is essentially one for the short and intermediate term, say, the next five years. It posits that, after the bottom is reached (which, it is hoped, will happen soon), the upturn will be slow because of balance sheet considerations. Offsetting that is the rapid increase in aggregate demand that can be expected in some countries where large fiscal stimuli have been applied and are beginning to take effect. Whether balance sheet or aggregate demand effects will dominate is difficult to say, but even if balance sheet effects do constitute a drag on the rate of economic growth in the industrial countries, the outlook will be for a gradual acceleration of the growth rate over time. Some countries with very high savings rates, most notably China, will almost certainly realize that it is in their self-interest to encourage the expansion of domestic consumption relative to gross domestic product (GDP), which would and should take up some of the slack as those in the previously low-saving countries save more.

The second concern arises from the observation that earlier growth was spurred in significant part by the increase in indebtedness of households and corporations in the industrial countries. While that is certainly true, it is also true that China, India, and others grew rapidly while accumulating assets, and as they continue to grow rapidly in the future, they will become more important as a source of international demand. While some of the industrial countries may grow more slowly than they did in the middle of this decade, other countries may accelerate their growth, and there seems little basis for pessimism over the medium or longer term about world growth prospects because much of that growth was financed by unsustainable borrowing. It will, of course, require a rebalancing of sources of demand, with China and other surplus countries spurring their own domestic consumption and reducing their current account surpluses, but that is in those countries’ interests and probably would have happened even without the crisis.

The third consideration—crowding out by the public sector—is the one to which most attention needs to be given. The International Monetary Fund (IMF) estimates that the public debt of the industrial countries will increase by about 85 percentage points over the period between now and 2015. As recovery progresses, it is likely that higher real interest rates will be needed to induce investors to hold public debt. Moreover, the aging of populations in the industrial countries and some others is certain to place a heavier fiscal burden on governments, even beyond those resulting from the current crisis. Both intermediate and longer-run considerations suggest that real interest rates and the cost of capital will be higher than in earlier years.

Policy makers in developing countries will face scarcer capital, either in the form of reduced capital inflows altogether or in the form of higher real interest rates to induce the desired inflows. In terms of the economic environment going forward, higher real capital costs are likely to be a significant factor once the upturn has become entrenched.

On balance, the argument that world economic growth will inevitably be slower in the future than in the past is weak. With appropriate policies, it is quite possible that growth could resume, with fewer structural weaknesses than were present in the years of “global imbalances.”
Impact of the Crisis on Developing Countries to Date

Initial discussion of “delinking” emerging markets and other developing countries from the industrial countries, should they enter a recession, was predicated on the assumption that the recession would be confined largely to one or a few industrial countries and would have about the same severity and duration as earlier recessions. As both its magnitude and its breadth exceeded all predictions, the impact on developing countries and emerging markets has been greater than anticipated.

However, although there have been serious repercussions for developing countries and emerging markets, they have been far smaller than expected based on experience in earlier recessions. Not only did most countries experience rapid growth in the five years prior to the crisis, but they also managed to tame their fiscal policies to a considerable degree. As such, many entered 2008 in much better fiscal and monetary shape than they had been before entering earlier periods of downturns in world trade and capital flows.

One of the impacts on developing countries is the severe reduction in capital flows and credit from financial institutions in industrial countries. For some countries, the drying up of foreign credit has itself been a major source of difficulty, especially for those with large current account deficits.

But for most countries, the differences lie in other areas. To a first approximation, countries can be divided along two lines. On the one hand, countries can be classified as oil exporters, other commodity exporters, and exporters of manufactures. Within that classification, there are degrees of reliance on foreign trade. On the other hand, countries can be classified by their fiscal stance and level of indebtedness. Within that classification, some entered the current period in much better condition than others.

Countries worst off include those that both are heavily dependent on one or more primary commodities whose prices have fallen drastically and have relatively high fiscal deficits and high ratios of debt to GDP. Argentina is such a country: exports consist largely of primary commodities and constitute a sizable fraction of GDP. Simultaneously Argentina’s debt-to-GDP ratio is high, not even counting the unrestructured external debt still outstanding. Meanwhile, Argentina’s inflation rate was officially around 12 percent before the crisis, but economists regard 23–27 percent as a more realistic number, and the government was appropriating pension funds in order to finance its activities. Clearly, there is little scope in such countries for countercyclical fiscal policy (as there were deficits during good years), and the impact of commodity price declines could be large.

By contrast, Chile had a budget law requiring a surplus equal to 0.5 percent of GDP in the structural budget, had accumulated funds during the years of high copper prices, and had little debt. While still affected by the slowdown in economic activity and the sharp drop in copper prices, the Chilean government had more than enough fiscal space to adopt a stimulus package. In contrast, in most prior recessions, the Chilean government had been forced to undertake procyclical fiscal tightening.
because of its starting fiscal position. Most forecasts are for Chilean real GDP to be about the same in 2009 as in 2008.

Likewise, one can contrast other countries. India and China are both exporters of manufactures, but China’s exports are a much larger percentage of GDP than India’s. China has large reserves and huge current account surpluses. India, meanwhile, has a high (above 80 percent) debt-to-GDP ratio and a large fiscal deficit (above 5 percent in the last fiscal year). The country has sizable reserves, but very little fiscal space given the large fiscal deficit and high debt level. It is likely to be less able than China to undertake policies that sustain GDP growth, even though China’s dependence on manufactured imports is considerably greater.

One can classify countries along these lines, and in general the larger the initial fiscal deficits and debt are and the more dependent countries are on exports of primary commodities whose prices have fallen sharply, the greater has been the relative severity of the impact of the current economic slowdown.

Worse, for countries with high initial debt levels and fiscal deficits, the room for offsetting fiscal action is severely restricted. Indeed, in some countries fiscal consolidation may even prove stimulative, in that the crowding out effects of the fiscal situation are sufficiently serious so that reducing the deficit may increase domestic demand.

Few developing countries, however, have thus far been able to adopt structural balance rules of the type that Chile (and several industrial countries) have. One lesson for the future is the desirability of such rules as a means of greatly reducing the impact of slowdowns in the world economy. If real interest rates in the world economy rise as anticipated, such rules will become even more important in the future than they were in the past.

**Development Policies in Light of the Global Outlook**

The outlook for the international economy as a whole is for growth that is no faster, and possibly slower, than the growth realized in the past very successful years. It is for an international economy with higher real interest rates and, it is hoped, smaller global imbalances.

The lessons for development policy stem from these observations. Outward-oriented growth strategies and integration with the international economy will still bring benefits, although they may be somewhat smaller than in the past (although growth will still be more rapid than if countries attempt to reinstate tariff and other barriers to imports). And, as long as any slowdown in overall growth is not the result of, or accompanied by, increased protection in the global economy, the benefits to integration will be sizable. With the recession, there is a risk that some policy makers will increase protection, which in turn would lead to retaliation by others. Major increases in trade barriers would serve as a significant damper on world economic growth in the longer term, with the largest impact on those emerging markets dependent on trade and on poor countries whose best hopes for more rapid growth lie in opening up their economies.
But if real interest rates are higher, and capital is therefore more expensive, a key implication is that the efficient use of unskilled labor, and efforts to increase the productivity of labor, especially by investments in health and education, will become more important than ever. As capital becomes more expensive, it will generally raise the optimal labor-capital ratio for goods production, which should be to the advantage of countries with abundant unskilled labor and to the advantage of those countries accumulating productive human capital.

As industrial countries go through the demographic transition and have a much higher fraction of older people in their population, investments in developing countries with relatively abundant supplies of unskilled and skilled labor will become increasingly attractive. But investment will flow only to countries where returns are attractive, and they will be attractive primarily in countries that have business-friendly investment climates (including, especially, open trade regimes), have invested in the health and education of their population, and have flexible labor markets.

For countries meeting those criteria, the opportunities even in a slower-growth world will be substantial. It will be to the advantage of savers in developed countries (where scarce labor will dampen returns to capital) to invest abroad where labor is more abundant, and it will be in the interest of developing countries and their workers to encourage these investments.

These longer-term goals cannot, however, be achieved overnight. Especially when it comes to education and health, investments in human capital have long lead times. Countries where those investments are considerable and growing will be in the best position to take advantage of the postcrisis, longer-term economic landscape.

But, as noted, even as countries adapt their policy framework to sustain growth in the new circumstances, the growth of the international economy will affect them. The higher that growth is, the greater the payoff will be for appropriate policies. For that reason, it is incumbent on leaders in developing countries to participate in international forums far more actively than they have in the past, especially when issues surrounding the open multilateral trading system are under review. Korea, the host country for this conference, has done well in making the case for keeping trade open and should be supported by all developing (and developed) countries.

To date, developing countries have to a large extent been free riders on the open multilateral trading system, benefiting from lower tariffs and other trade barriers negotiated at the World Trade Organization primarily among developed countries, while unilaterally undertaking trade liberalization without binding tariffs. Some of the dangers of this practice for the international economy were seen in late 2008 and early 2009, as some countries took advantage of the space between bound and actual tariffs to raise their tariffs against imports. Much of the protection that resulted was by one developing country against another.

But even more important, protectionism not only hurts the country undertaking it but also invites retaliation, especially in difficult economic conditions, such as the present. Strong political leadership is needed to resist protectionism during the crisis. Trade barriers are politically much easier to raise than to lower. Raising a trade
barrier meets the approval of people within the industry, and those who are adversely affected (as consumers pay more or get lower quality for some goods, as businesses pay more for their inputs and are therefore disadvantaged in competing in export markets, and as overall growth slows) usually do not recognize the extent to which they are disadvantaged, certainly not immediately.

As the Korean experience amply demonstrates, the opportunity to use international markets is vitally important for rapid economic growth, especially in the early stages of development. No country has sustained rapid growth behind high and rising walls of protection, and no developing country has sustained a successful long-term development strategy without opening up to international markets.

The reasons why liberal trade and trade opening are such a powerful stimulus to growth are not entirely understood. Some of the benefits, such as increased competition for domestic producers and access to international best practice, clearly depend more on openness than they do on the export opportunities that increase with open trade strategies. Even slower growth of the international economy would not deprive countries with outward-oriented trade strategies.

Protecting and enhancing the open multilateral trading system ought to be a top priority for the leaders of developing countries in international forums. Success would raise global growth prospects, while simultaneously supporting the economies of those developing countries.

A leaner world economy, with less scope for policy mistakes, can be expected in the future. The graying of developed-country populations can prove advantageous for developing countries that adapt their own policies appropriately. The falling labor-to-capital ratio in developed countries can lead to a shift in comparative advantage to even more labor-intensive activities in developing countries. In an appropriate economic environment, which includes the rule of law, a good commercial code, enforcement of property rights, attention to the productivity of labor through education and other social expenditures, and a flexible labor market, the rewards in terms of growth and rising living standards may be even greater than in the past. The problems arising from fiscal crowding out and the demographic shift in developed countries provide an opportunity for developing countries to use the international economy to their great benefit.

References


My task is to speak on the global crisis and to address the issue of whether it is over or in what sense it is not over. I should make it very clear that I speak only for myself today and that none of my previous or current employers or associated organizations would in any way endorse what I am saying; these include the Congressional Budget Office, where I am on the Panel of Economic Advisers. I draw heavily on work with my co-authors, particularly James Kwak, who is, as Mr. Shin said, co-founder of my Web site BaselineScenario.com.

The quick answer to the question, is the world crisis over yet? is a reassuring yes. Confidence is returning to financial markets. We have not yet seen a full stabilization, particularly with regard to unemployment, but we have seen substantial improvement. Some stability is returning to financial markets, and that is very good news. Perhaps we should just relax and take the rest of the afternoon off. At the same time, in a very real and important sense, the crisis is not over. Not only is it not over, but the underlying problems that brought us into this situation are, in my view, still with us and have actually gotten worse.

In this address, I will try to explain the role in which an oversized financial sector in the United States but also in other parts of the industrial world, particularly but not exclusively in Western Europe, has contributed to the buildup of vulnerabilities over the past six months to two years. And because of the way in which the short-term problems are being addressed—the U.S. government is throwing a massive amount of money and other resources at the U.S. banking system—these vulnerabilities are getting worse rather than better.

I am an optimist, and nothing about my message should be considered dark or pessimistic. There is going to be a 5- to 10-year reform process in the United States...
and perhaps in other countries, and eventually real reform will come. The financial sector will be substantially changed in the United States, and that will be a good thing for all of us. But the United States has already missed the opportunity to act immediately—by which I mean this year—and the costs of missing that opportunity are enormous. The costs are significant for American citizens and for the American economy, but they are potentially much more traumatic for the rest of the world, including East Asia and the Republic of Korea.

You can take two views of the crisis. One is the official view, which you hear from U.S. government officials; it is the G-20 consensus. It is expressed most articulately in the United States by Larry Summers, who is, of course, the former secretary of the Treasury, the current head of the National Economic Council in the White House, and President Obama’s right-hand person for all things economic. Larry Summers says that we have experienced a rare and unfortunate accident. It only happens about once a century, in his view, although, by his own account, we have experienced about 10 major financial crises in the past 20 years. Nevertheless, the crises that happen once a century are very complex and have to be sorted out by the experts. I will come back to that point.

Such a crisis needs to be counteracted with a massive, absolutely huge, macroeconomic policy response, which is unprecedented for peacetime. The last time the United States or any other country did something like this was at a time of war. I believe that the ratio of U.S. debt to gross domestic product (GDP) is going to double as a direct result of this crisis and the measures taken to counteract it. This is my personal opinion, which is why I am leaving the Congressional Budget Office out of this. I believe that U.S. debt will grow from about 41 percent of GDP to around 80 percent. That is not necessarily a catastrophe. There is a view that the U.S. can afford this level of debt, but it is a large increase in debt for no good purpose. In addition to this enormous macroeconomic response, the great deal of regulatory forbearance for the banks, and the substantial amount of cash being pumped into the banking system (perhaps the full extent of which we do not yet know), there is also a regulatory response, which is tiny.

This is where U.S. policy currently stands, and I track this on a day-to-day basis. I present my arguments and contrast them with the official view through my Web site. I testify before Congress quite frequently. I run a Web site for the Washington Post that tracks congressional hearings and the debate around economics. I am quite familiar with the details. The proposed regulatory changes do not go very far. The official view is that the crisis is under control: “Don’t worry. Go back to your homes. Be happy.”

The second view is somewhat different. It says that a political and economic structural change has taken place in the United States and other countries since the 1980s. Essentially a U.S.-led change created global vulnerability as a result of the destabilizing power of the financial sector. This experience is not at all unusual. It is a standard part of U.S. history and global history, although we have not seen it for the past 50 years in the United States. In some sense, we have returned to previous experiences in the United States and in other parts of the world, but we have returned to them in
a particular kind of globalized economy in which the United States is playing a particular kind of role.

Many measures have been taken to stabilize the system. You can track them on my Web site. The approach is big, and it is expensive; it works in the sense that if the government puts huge unconditional and potentially unlimited subsidies into the banking system—if it has the fiscal capacity to do that—the system is going to stabilize. The government has reduced the probability of bank runs and bankruptcy, but there is still a level of risk around a couple of major U.S. banks. This is extraordinary given the support that has been provided.

As a result, the government has created more job security for insiders. Very few executives—the leaders of the major banks who destroyed the value of their companies and wiped out most of their shareholder value—have lost their jobs. It is a terrific time to be running a bank. The official approach has also helped stock investors, at least for a while. Some confidence has returned to the markets, although it might be temporary. What is wrong with that? What is wrong with pumping massive amounts of subsidies into a troubled sector? The official view is that an accident occurred and we should feel sorry for these poor, downtrodden banks and bankers.

The problem is that in the United States and around the world, powerful groups tend to rise, particularly on the back of long economic booms, and they tend to take over the political system. That is very bad for society as a whole.

The United States has gone through exactly what we are facing now at least three times in its history. In the 1830s there was a confrontation between then-president Andrew Jackson and the Second Bank of the United States, which was the financial powerhouse of its day. It was a nasty confrontation that lasted at least five years and resulted in the destruction of the bank, with plenty of collateral damage. The executive power—the power of the presidency—prevailed in those circumstances. At the end of the nineteenth century, we also experienced the rise of the big railroad trusts, which were also based on banks and on access to finance, so industrial power was interwoven with financial power in that episode. Teddy Roosevelt confronted this confluence beginning in about 1901, and, again, the fight was very difficult, protracted, and vicious, lasting close to five years before, once again, the political power of these industrial and financial groups was curtailed. And, of course, the Great Crash took place at the end of the 1920s and the beginning of the 1930s. Perhaps the most famous confrontation between political power and finance in the United States occurred around the so-called Pecora hearings, which exposed many of the wrongdoings of major financial firms, including the forerunner of what is now Citigroup and the forerunner of what is now JP Morgan-Chase. These hearings resulted in modern securities law and much of the legal foundation for the postwar regulation in the United States that worked quite well, at least through the 1980s.

The highly regulated banks of the 1940s to 1970s in the United States should not be seen necessarily as the natural state of things. They were part of a historical cycle of deregulation and reregulation that has characterized 200-plus years of American financial history. It is a cycle seen in many other parts of the world.
There is, however, a difference between now and the end of the nineteenth century. At the end of the nineteenth century, the United States had large railroad trusts with big banks behind them, and the problem was one of monopoly power. Some textbooks state that the Sherman Antitrust Act of 1890 created a clear legal foundation for breaking up these trusts, but that is a simplification. In fact, Teddy Roosevelt made a very deliberate and difficult political decision to confront the major trusts, particularly those backed by the House of Morgan (JP Morgan). This decision led to the political struggle and reforms of the early twentieth century.

The difference between that situation and today is that the large banks, again, have extraordinary political influence in the United States and elsewhere. Part of this influence is based on what I consider to be largely false claims of financial innovation, which involved consumers overpaying for products in a way that is quite close to the overpaying that occurred as a result of the monopoly power surrounding railroad fares, which was the big deal at the end of the nineteenth century. In addition, a much worse problem is that the banks have figured out, perhaps inadvertently, how to extract rents from the state. This is not a troubled emerging market or a developing country. This is the United States. The most sophisticated, advanced financial system in the world has come to the point where banks are extracting enormous rents directly from the state and are forcing the government into a position where officials feel they have no alternative but to engage in this enormous increase, unprecedented in peacetime, in the public debt.

This is not what is commonly referred to in the United States as regulatory capture. There are some features of regulatory capture, but this is essentially state capture. This is the power of the banking system to cause damage in order to transfer rents from the state to itself. This power has not been seen in the United States for 50-plus years and has not been seen in industrial countries for a long time. It is not corruption, such as seen in Indonesia at the time of President Suharto or in the United States in the nineteenth century. The United States is not exempt from corruption, but it is on a small scale. It is not the kind of political connections that, for example, characterized the functioning of Malaysia under Prime Minister Mahathir or, again, that were evident in the United States in some previous historical episodes.

The United States has the world’s most advanced oligarchy. Oligarchy is an uncomfortable word for Americans, and I do not use it lightly. It is a word that I and others introduced into the popular economic lexicon earlier this year with some trepidation. But it has caught on, and it does convey the essence of the idea, as defined by Aristotle, that political power arises from economic power. The banking sector has converted enormous economic power into political power, particularly through what we call a form of cultural capital. It has persuaded people that letting finance run unfettered in the 1990s and 2000s was good for the economy. What was good for Wall Street was great for the U.S. economy and great for the global economy, although there was some hesitation about that after the 1997–98 Asian currency crisis. Part of this power comes through campaign contributions. Both parties—but particularly right now the Democratic Party—rely heavily on
donations from Wall Street, for example, to fund congressional races and support the general fund.

The problem is more damaging and more dangerous than corruption or political connections or straight campaign contributions. It is a form of intellectual capture. It is a form of persuading people that there is a genius of finance and that the genius has somehow transformed the nature of productivity and the nature of lending money; therefore, banks should be allowed to do whatever they want. That is wrong. It is incredibly dangerous. It has got us into this very difficult global economic situation, and it is not being fixed.

In an article published recently in the *Atlantic Monthly* (Johnson 2009), I explain how this story, this political economy, what happened in the United States, is not unique to the United States. It is a particular manifestation of issues and developments seen much more broadly in other episodes around the world. As a result of deregulation in the 1980s, major finance players gained rising economic power in the United States. This money was put back into politics and into buying intellectual influence. The bandwagon of banks was immensely alluring to many, and this helped to build arguments for further deregulation and for running relatively easy monetary policy. The movement in this direction was greatly helped by the arrival of “new technologies.” One was the availability of emerging markets that were potentially open to capital flows, and this is something that Jagdish Bhagwati (1998) has written about eloquently in an influential article called “The Capital Myth.” He said that this was very dangerous, and he was absolutely right. What he did not fully realize—what none of us realized at the time—was that the biggest danger for emerging markets and for the world was the effect of unfettered finance on the United States. This, of course, was particularly the case once derivatives became more readily tradable and once the falling costs of computing power made it easier for more people to enter those markets. All of this created more economic power for the big banks.

The story that I am telling you—a boom based on initial economic advantage turns that advantage into political privilege, tilting the playing field—is, of course, a standard emerging-market boom. It is Russia or Ukraine or Brazil in various episodes. I am not saying that it is Korea. The boom brings five or 15 years of good growth, but it entails taking a lot of risks with borrowed money. This occurs, in part, because investors feel that the financial system has a lot of political power and that the state will stand behind it. The overexpansion creates vulnerability to shocks, and, of course, this comes to a grinding halt through some combination of a currency crisis, banking crisis, and fiscal crisis. The particular form that this took in the United States was, of course, the enormous deregulation of the 1990s.

The irony is that the current economic team in charge of stabilizing the U.S. economy is the same group of people who were in charge of these policies during the 1990s. I do not say that to disparage them. It is merely a statement of fact. I have no problem at all with them changing their minds. Larry Summers and Tim Geithner and others have changed their minds. That is what they say, and I believe them, but these incredibly influential, smart, experienced people bought into this myth of the
genius of finance in the 1990s, and the resulting policies mattered. This is what the United States did to itself and, by implication, to the rest of the world.

What breaks this kind of crisis? What breaks a crisis that has been built around this concentration of political and economic power? Some of the powerful groups—we can call them oligarchs; we can call them something else—fail or lose their businesses. In an emerging-market context, there typically are not enough resources to bail them all out. The International Monetary Fund (IMF) may get involved, which comes with mixed publicity, but its involvement often is helpful. We can talk about that, particularly when the IMF diagnosis focuses on the oligarchs and on strategies to contain the rent seeking of the oligarchs. Of course, the United States is different. For one thing, the IMF will not be called in. I can assure you of that. Second, the dollar is a reserve currency, and it has enormous fiscal capacity; it is the borrower of first resort or the place where investors park their money when the world hits difficult times. So the United States has enough resources to bail out much of big finance, though not all of it, and this resembles the kinds of bailouts we saw in Japan in the 1990s, particularly if the fiscal stimulus effects are included. The government has enough resources to create a temporary system of rents that can be distributed to these banks going forward.

Have the bankers won? They have, at least in the short term. There will be an economic recovery. An economy the size of the United States and with a degree of flexibility and diversification does recover, and the crisis strengthens those bankers who remain. Jamie Dimon, who is the chief executive officer of JP Morgan-Chase, which is now the largest bank by far in the United States and remains independent of state supervision, said to his annual meeting of shareholders a couple of weeks ago that they had probably seen their best year ever. That was 2008. That was the global financial crisis from the perspective of JP Morgan. The top three U.S. banks now control about 30 percent of all deposits, up from about 20 percent two years ago and up from around 10 percent perhaps five or six years ago.

Over the longer term, I am not so sure the bankers have won. In fact, I think they probably have not, in part because there will be a lot of overgrazing. The bankers cannot hold back. They cannot refrain from paying themselves large amounts of money or from engaging in egregious compensation schemes and other things that will greatly alienate and antagonize people. There is much more public scrutiny of excess. People are watching for errors, and they will find them. This kind of growth is unlikely to prove sustainable. A banking system like this is going to be volatile. Other powerful groups are very unhappy. The calls for reform are getting louder.

The official view is that it is just the populace, me, and some of my friends versus the bankers and the people who run the country. We are tilting at windmills. But this is not true. I talk to people around the United States all the time about this. I engage with people through the Internet and through the regular media, and I can tell you that most people in the United States who have thought about this for more than five minutes are extremely upset by the current situation, not just by what occurred before 2007, not just by the way things were mishandled in 2000 to 2008. They are upset about what is happening right now.
This is particularly true among the people who run small banks. Small banks are allowed to fail. The Federal Deposit Insurance Corporation (FDIC) has taken over at least 30 banks so far this year. Perhaps they have taken over even more. Small banks fail with some regularity. The United States is world class at liquidating small banks, and small banks are very upset with these large banks that have been deemed too big to fail. Venture capital is very antagonized by this. I spoke to a large technology conference of venture capitalists and entrepreneurs in March of 2009, and I said to them what I just said to you. They did not sit there quietly, as you are sitting there quietly. They were very agitated even before I arrived on the scene. Private equity, which is an important political power within finance in the United States, is deeply antagonized by the way the large commercial investment banks have behaved. Private equity could go either way. They are fairly flexible in political terms, but I see potential for them to join, in some form, a coalition of people who would like to change the system, not just the financial system. Most people are beginning to be or are already quite upset about this situation.

Why am I making all this fuss? Why can’t we just hold our noses and put more money into the financial sector and continue with what we have been doing? I see a couple of problems with that. First of all, financial innovation, or so-called innovation, has been very harmful to consumers. Most financial innovation since the 1970s has not been like nonfinancial innovation. I believe that some form of greater consumer protection is coming to the United States, probably in the form of a new agency that will focus on protecting consumers against dangerous financial products. I do not think that will be an easy or a quick fix, but it is going to happen. This is going to limit the ability of the financial sector to take advantage of consumers. At the same time, banks have found that they are too big to fail, and this moral hazard is going to affect their behavior. JP Morgan-Chase may well have been a well-managed bank in this crisis, and the risk culture of JP Morgan is applauded by many. The leadership of Jamie Dimon is particularly applauded by people in power, and I have no problem at all with the individual. I am sure that he is brilliant and far-seeing and cautious, but you cannot build a banking system around a couple of guys who are sensible and did not get carried away in a boom. You have to have a system. The system of incentives is, very clearly, that if you take large enough risks and fail in a sufficiently synchronized manner, you will be bailed out. The government has blinked; to my mind and to the mind of almost all the independent observers and analysts who have looked at this, nothing in the government’s reform package is going to change these incentives substantially.

The real reason that reflated finance is not going to save the day is that finance is already very large. It is unsustainably large, and its share of corporate profits and of total compensation is extraordinarily high and hard to sustain. In addition, the share of talent that goes into this sector is breathtaking. Claudia Goldin and Larry Katz have a very interesting project, titled “Harvard and Beyond,” looking at what has happened to Harvard graduates. As an MIT professor, I am happy to pick on Harvard, although I do so only because we have the data. Since the beginning of their data, Harvard graduates have gone increasingly into finance. Basically, before 1990
between 5 and 15 percent of Harvard graduates went into finance. Since 1990, the percentage has doubled to perhaps 20 to 25 percent. The education elites have gone increasingly into finance. (See Johnson and Kwak 2010 for a longer discussion and the full reference.)

The second issue pertains to profits in the U.S. financial sector. Larry Summers now says that when we saw financial sector profits as a share of total domestic profits hit 40 percent in 2003, we should have taken that as a warning sign. I agree with that. It is very hard to imagine financial sector profits higher than they were in the 1980s.

The third issue is U.S. financial sector compensation. Up until the 1980s, compensation was roughly the same in the financial sector as in all private industries. You got paid the same in banking and finance through 1980 as you did in other sectors. The compensation of financial engineers and nonfinancial engineers was roughly comparable, for example. That changed in the 1980s. If you went to work in finance, you increasingly got a premium. Ariell Reshef and Thomas Philippon have done some very nice work on the origins of this (Philippon and Reshef 2009). They argue, I think persuasively, that compensation in the financial sector was high prior to the 1930s, but the 1930s regulations of the Glass-Steagall Act, the Securities Act, the Securities Exchange Act, and so on pushed down compensation in the banking system. Banking became boring. This is Paul Krugman’s phrase, and it is a very good phrase. Banking was boring and not well compensated in the 1950s, 1960s, and 1970s. In the 1980s, deregulation removed interest rate ceilings, among other things. The first wave of deregulation occurred from 1980 to 1984, and a second occurred in the 1990s under the supervision of Robert Rubin, Larry Summers, and other people who are now back at the Treasury and the White House. Compensation rose, and, of course, this is part of political power and the origin of campaign contributions. This is the origin of cultural capital. This is why finance was sexy, to use a technical term. People wanted to go into finance. It seemed to have won and to have prevailed. It pulled a large chunk of the educational elite, including professors at business schools, into the financial sector. It did so in this remarkable, hard-to-sustain pattern. Thomas Philippon, for example, who offers an articulate defense of the financial sector, and James Surowiecki, who is a popular writer in The New Yorker, argue that compensation will decline as a percentage of GDP, to reach levels in the mid-1990s. (See Johnson and Kwak 2010 for a full discussion.) Finance plus insurance as a percentage of GDP has basically tripled in the United States since the end of World War II. They say that finance will give up 1 or maybe 2 percentage points of GDP. That is a big contraction. That is not a reflated financial bubble. And that opinion comes from the defenders of finance.

Something else is strange about this picture. Let us look at the well-known story of agriculture as a percentage of GDP. Technological innovation means that a staple good—food and everything related to agriculture—is produced using far fewer resources, particularly less labor, and is shrinking as a percentage of GDP. If finance is so wonderful and so innovative and so productivity enhancing, why does it go up as a percentage of GDP so consistently?
I have picked on the United States rather a lot, in part to be provocative and in part because the crisis started in the United States. The rise of finance and the victory of finance are an American phenomenon, but, unfortunately, the intellectual capture by finance is not peculiar to the United States. If this were just about the United States, the United States would have a big problem. The world economy would have a serious issue, but we would not have a global crisis. We have a global crisis because Western Europe bought into this story and the myth of the financial sector perhaps even more than the United States did. The United States has banks that are too big to fail—at least that is what the Federal Reserve thinks and what the Treasury thinks. Western Europe also has banks that are too big to fail, but it may also have banks that are too big to rescue. Western European banks are much larger than U.S. banks as a percentage of the economy. The largest banks in the United States peaked at a total balance sheet of 20 percent of GDP. The largest bank in the United Kingdom—the Royal Bank of Scotland—had a balance sheet at its peak of more than 100 percent of GDP. That is just one bank. That is just one failed bank that has been bailed out. Finance as a share of GDP has increased consistently across developed countries. This observation is based on data used by the Organisation for Economic Co-operation and Development (OECD) and the Bank for International Settlements. These data suffer from definitional problems, but they show the general pattern of an increase in the contribution of finance to GDP.

The OECD also calculates something called the excess credit level, which has its own methodological issues, but which conveys the deviation from the long-term trend of domestic bank lending to the private nonfinancial sectors as a share of GDP. Once again, something happened in the late 1990s: excess credit increased in the United States and in the euro area. Iceland, of course, stands out as extraordinarily irresponsible, with total bank assets around 11 or 12 times GDP; according to the same data, total bank assets are six times GDP in the United Kingdom and eight times GDP in Switzerland. It does not matter how you cut the data, the story is the same. The United States has had a big increase in financial sector problems and in the vulnerabilities that have built up because of the financial sector, but it has not had a big increase in bank assets as a percentage of GDP. That has happened in Switzerland, in the United Kingdom, in Ireland, and in some other Western European countries.

Who is responsible for the global crisis? Of course, the answer is complex, and there are many proximal causes. We can talk about housing. We can talk about the overexpansion of credit. We can talk about excessive risk taking by financial institutions. All of those explanations are, in some sense, right, but we need to look at the deeper cause. We can call it a meta bubble, or we can call it the new oligarchs. It is the rise of the financial sector in the United States and in Europe since 1980.

Should we care that finance has become so big that it has changed the nature of politics? Well, institutions matter. Weaker institutions do not prevent growth, but they make it very hard to sustain growth. They make it easy to enter a Latin American pattern of growth, boom, bust, and lost decades. Weak institutions give rise to more frequent crises, more severe crises, and derailed growth.
I am talking here about the United States, which is not on the periphery of the world’s financial system. It is right in the middle. Addressing this problem will dominate all other considerations of whether or not the world can grow and whether or not the World Bank can help to reduce poverty around the world. This swamps everything. If this goes wrong, if we do not disengage from finance, if we do not find a way to disintermediate, we can say goodbye to moderation and hello to instability. The costs will be felt in the United States in the form of higher taxes, slower growth, and perhaps more inequality. The costs outside the United States will be much more damaging. They will be felt in higher poverty, shorter life expectancy, and perhaps higher inflation. Maybe this is the great fail-safe, the escape hatch. There is plenty to suggest that the U.S. strategy is moving in this direction, although, of course, the officials deny this vehemently. However, I think that it is true, and the financial markets think so do.

If you are moving into a much more emerging-market structure of the economy, then you can have a lot of inflation. The standard macroeconomic models say, “There will be no inflation until 2010 or 2011.” Standard macroeconomics also said that we could not have the kind of global crisis—with the collapse of output and the massive decline in world trade—that we have had over the past 18 months. So it is not clear that those standard models are adequate to explain the situation.

Many people do not agree with my perspective. David Brooks, a well-known newspaper columnist in the United States, has offered the most interesting counterarguments. He has said, “Simon Johnson is wrong. The banks are not super smart or conspiring.” I do not say that the banks are conspiring or that they are super smart. I do say that we built this system and that bankers have taken over the political system and acquired a tremendous amount of power. In fact, they are really stupid. If the banks are so stupid that they cannot manage the risks that they take on, then why do we allow banks that are too big to fail to remain? The economic solution is very easy, but the political solution is very hard. Breaking up the trusts in 1901 was very hard politically, too. But we need banks that are small enough to fail. We need to break up the big banks in the United States and elsewhere. If they can fail, that is fine. If they cannot fail, if the government has to come in and bail them out one way or another, that is very dangerous.

Another view says that the experts who built the system are needed to solve the problems. That may be true, but do you really want to leave a system in place where the next time the equivalent of AIG’s financial products fails you have to keep these people on with million-dollar retention bonuses because they are the only guys who know how to unwind those trades? That is not acceptable to me. A consumer protection agency is going to help, but only a bit. The financial sector can get itself into trouble in many ways, even without taking advantage of unsuspecting consumers.

The strongest pushback I get from this view is from people who say that I misunderstand what the Obama administration is doing. They say that it is not captured by an oligarchy and that it will implement reform. The most extraordinary and alarming argument that I hear from a few people is that we should not make this argument at all; we should not study the nature and the workings of political con-
nections in the United States because nothing good will come of it. I have heard that we technocrats should stick together and stick with big finance. I do not agree. I do not think that it is good economics. I do not think that it is good politics. I do not think that it is good for the U.S. economy. I do not think that it is good for the world economy.

I am a professor of entrepreneurship at MIT. I spend a lot of time talking to entrepreneurs and venture capitalists. I am not a far-left radical. I am not a far-right radical either. I am a complete centrist, and I work with entrepreneurs in the United States and around the world. My view, and the view of the venture capital sector and the view of the powerful people in that sector, is that crazy investment banking—speculative financial markets—is in no way essential, helpful, or constructive at this point to what is needed for innovation and growth in the United States and more broadly. The major risk to innovation and growth in the United States and elsewhere in the world has always come from a rent-seeking sector. The name of this sector changes from time to time. It is not the railroad barons anymore. It is big finance. We are in a cycle, and the cycle is one of problem and reform, but reform takes a while. Break up the big banks, make them small enough to fail, preferably do it sooner rather than later, or we will all face the consequences.

In conclusion, my simple, straightforward, and provocative message is that there has been a rise of finance capital and political power in the United States since the 1980s. This has repeated a historical pattern seen in the United States before. It is common in the long booms that have been frequent in U.S. history. It is also very familiar in emerging markets. There are parallels or tagalongs in other industrial countries, including much of Western Europe, and it is important to keep in mind that having a crisis, by itself, solves nothing. The surviving oligarchs become stronger. Their intellectual market share, if you like, goes up.

Will the twenty-first century turn out to be a great deal like the end of the nineteenth century and the beginning of the twentieth century, with a big argument about what needs to be done to deconcentrate economic and political power in the United States? I think that the answer is yes. I think that we will experience a recovery and that the world economy will stabilize and turn around. Growth may return to reasonable levels quite quickly, depending on the ultimate extent of damage to people’s balance sheets and to consumer and investor confidence. But that is not the end of the crisis. That is not the end of the discussion. That is not the end of vulnerability. If recovery is just around the corner, so is another crisis, which could cost 40 percent of GDP or more in the United States in terms of additional debt. Perhaps the level of U.S. debt will go from 80 to 120 percent of GDP. Perhaps we can afford that. It seems like a waste to me, and it seems as though such levels would prevent President Obama from undertaking many sensible initiatives, but that may be what we are facing.

The point I would leave you with is that whatever happens in the United States, whatever the costs in the United States, whatever the damage done by the United States to itself and its own people with this kind of rent-seeking, overly powerful financial sector, the damage to the rest of the world, to ordinary people trying to
make their way in the modern globalized world, will be much greater. Unless we deal with this problem soon and effectively, the consequences could be dire.

References


Industrial Policy and Development
Industrial Policy and Development: A Political Economy Perspective

JAMES A. ROBINSON

This paper discusses the role of industrial policy in development, making five arguments. First, from a theoretical point of view, there are good grounds for believing that industrial policy can play an important role in promoting development. Second, there certainly are examples where industrial policy has played this role. Third, for every such example, there are others where industrial policy has been a failure and may even have impeded development. Fourth, the difference between these second and third cases rests in the politics of policy. Industrial policy has been successful when those with political power who have implemented the policy have either directly wished for industrialization to succeed or been forced to act in this way by the incentives generated by political institutions or the wider environment. These arguments imply that we need to stop thinking of normative industrial policy and instead begin to develop a satisfactory positive approach if we are ever to help poor countries to industrialize.

In this paper I discuss the role of industrial policy in development. I make five main arguments. First, from a theoretical point of view there are good grounds for believing that industrial policy can play an important role in promoting development. Second, there certainly are examples where industrial policy has played this role. Third, for every such example there are others where industrial policy has been a failure and may even have impeded development (though the counterfactual is complicated). Fourth, the difference between these second and third cases rests in the politics of policy. Industrial policy has been successful when those with political power who have

James A. Robinson is with the Department of Government and the Institute for Quantitative Social Science at Harvard University. The author thanks Alan Gelb for his indispensable help with the presentation and an anonymous referee for comments.

Annual World Bank Conference on Development Economics 2010, Global
© 2011 The International Bank for Reconstruction and Development / The World Bank
implemented the policy have either directly wished for industrialization to succeed or been forced to act in this way by the incentives generated by political institutions.

My fifth point, which is derived from the first four, is that economists and international institutions have to change the way they think about industrial policy. To promote industrialization in a society, we need a positive theory of policy. I use the term “political equilibrium” to refer to the concatenation of political forces that determines this policy outcome. It is the political equilibrium of a society that leads to particular policy choices. To give policy advice that would foster industry, one has to understand this political equilibrium and either attempt to change it or work within the environment it generates. This is a very different way of thinking about what industrial policy means.

Unfortunately, as things stand, while we have a good normative theory of industrial policy, we have a woefully inadequate positive theory, which can help to explain why industrial policy was adopted and apparently so successful in Taiwan, China, for example, and such a disaster in Ghana. It is toward building such a theory that research should focus, and without it, I argue, advocating industrial policy in the traditional sense as a solution to the problems of poor countries involves a large leap of faith.

Before delving into the consequences of industrial policy and my arguments in more detail, it is important to have some sort of definition of industrial policy. I take industrial policy to mean that the government deliberately attempts to promote industry. Nevertheless, there are many ways in which this can be done and many things that can count as industrial policy. These include tariff and trade policy (protection), tax relief, subsidies of various forms, export processing zones, and state ownership of industry. The way that Joseph Stalin promoted industrialization in the Soviet Union in the 1930s was completely different from the way that General Park Chung Hee did so in the Republic of Korea in the 1960s. Moreover, any of these policies may have promoted industrialization inadvertently, rather than deliberately. This is not an issue in the Soviet or Korean case, since the governments overtly committed themselves to industrialization. In other cases, however, it is not obvious whether or not industrialization occurred intentionally. One example would be the Calico Acts passed by the British Parliament in 1701 and 1721 with subsequent amendments. These acts raised prohibitive tariffs on cotton goods imported into Britain from India and even banned the wearing of garments made out of Indian fabrics (calicos). In fact, it was not until 1774 that the wearing of all cotton cloth was legal. By this time, of course, a vibrant British cotton industry had emerged. The traditional interpretation of these reforms rested on the notion that eighteenth-century British governments were laboring under the doctrine of “mercantilism”—an incorrect theory of how the economy worked. According to this view it took a revolution in ideas, induced by Adam Smith, among others, to lead to the final repeal of these acts in 1774. A more positive explanation for the introduction of these reforms was that they were advocated by the English wool and linen industries, which were suffering from Indian competition (Mokyr 1999, 50–51). Neither view suggests that there is any connection between the Calico Acts and industrialization. However, the Calico Acts, in conjunction with the Statute of Monopolies of 1623, which made it more or less impossible to establish domestic
monopolies in Britain, allowed the initially uncompetitive British cotton industry to develop without facing international competition. Although we do not know what would have happened in the absence of protection, it is obvious that the Calico Acts raised the return to investing in the cotton industry and may have played an important role in stimulating investment in the industry, which sparked the British industrial revolution.

Were the Calico Acts an industrial policy? If either of the first two views was correct, I would argue no: the stimulation of industrialization was an unintended byproduct. Nevertheless, neither of these views provides a compelling understanding of economic policies after the Glorious Revolution in 1688. In fact, as Pincus (2009) convincingly shows, the Glorious Revolution was led by a Whig coalition that quite definitely and explicitly had the object of stimulating “manufactures”—in other words, industrializing. To this end the Whigs started the Bank of England, facilitated the development of the transportation sector via canals and turnpike roads, reorganized the tax system, and changed commercial policy. In fact, the Calico Acts were part of a vector of policies that constitute one of the world’s most successful and most consequential industrial policies (the “mother of all industrial policies”).

There are quite a few existing approaches to the role of industrial policy in the development literature. Early work during the 1940s and 1950s, by Rosenstein-Rodan, Myrdahl, Nurkse, Hirschman, and others, associated development with industrialization but argued that, due to various types of market failures, one could not expect this to happen automatically in poor countries. Hence, an important role for the government was to stimulate industrialization with an industrial policy. A particularly important subset of these ideas was due to the work of Singer and Prebisch emphasizing dynamic comparative advantage and the need to close the economy for some period to develop an internationally competitive industrial sector.

These ideas were part of mainstream development economics until they came under sustained attack from scholars with a public choice bent in the late 1970s and 1980s. Scholars such as Anne Krueger (1993) and Deepak Lal (1983), anticipated to a large extent by Peter Bauer, argued that industrial policy had not worked and indeed could not work because government failures were always worse than market failures. One should forget about industrial policy or, for that matter, any other policy intervention to solve problems of development and instead focus on creating free markets and a night watchman state. This literature was certainly correct in pointing to some very unsuccessful instances of industrial policy in developing countries. However, it was selective in its focus. Moreover, the theoretical argument that government failures are always worse than market failures seems more ideological than based on either theory or evidence.

Inevitably, therefore, this view did not remain convincing for long, even if it had a large impact on development agencies in the 1980s. The most damning evidence against it came from a series of important interpretations of the “East Asian miracle” economies by Johnson (1982), Amsden (1989), Wade (1990), and World Bank (1993). These works all put successful industrial policy at the heart of the postwar economic successes of Japan, Korea, and Taiwan, China, respectively. This research,
and much more like it, stood the public choice view on its head, more or less arguing that market failures were always worse than government failures and that industrial policy was a powerful tool to promote economic growth. This interpretation of the East Asian experience heavily influenced the famous World Bank report in 1993.

Although the evidence in these studies is compelling, the evidence on unsuccessful industrial policy is equally compelling. This suggests that neither extreme view is correct. Industrial policy can sometimes work, but sometimes not. What distinguishes these cases? An obvious difference is that the type of industrial policy adopted in different cases was very different. In Latin America, for example, it came in the form of import substitution industrialization (ISI) with domestic markets closed to international competition. In Korea and Taiwan, China, the model was instead based on exports, with incentives created to induce the development of export industries (although it is also true that the domestic market was protected).

Why were such different strategies chosen to promote industry? As with the explanation for the Calico Acts, many scholars attribute this variation in policies to variations in ideas. Krueger (1993), for example, argues that Latin American countries were led by erroneous economic theories into adopting the ISI model. Possibly East Asian politicians had better (or maybe different and luckier) economic advisers than those who worked in Latin America. Ultimately, variation in the adoption and success of different industrial policies is explained by differences in the ideas and ideologies of different policy makers or their economists. Thus, Stalin had an industrial policy of a particular form because of his socialist ideology, while Mauritius had a successful export processing zone because Nobel Laureate Sir James Meade (1961) persuaded the governing Mauritius Labour Party that it was a good economic policy for the country.

An alternative approach to explaining variation in the success of industrial policy is set out in Rodrik (2007). Rodrik’s basic argument is that industrial policy is potentially very powerful, but one size does not fit all. For promotion to be successful, development industrial policy has to be tailored to the specific context or institutions of a country or, to use the terminology introduced by Hausmann, Rodrik, and Velasco (2007), has to be sensitive to the “binding constraints.” According to this view, different countries could adopt identical policies with very different results because they had different sets of market failures. Why would some countries adopt policies suited to their market failures and not others? The main reason concerns the binding constraints. Either Korea was very lucky in being able to understand this or it had (again) better economists who managed to understand this. Ghanaian industrial policy failed because some academic scribbler (actually Sir Arthur Lewis; see Lewis 1953) persuaded the government of Kwame Nkrumah to adopt an industrial policy that was not the right one, given Ghana’s circumstances.

In this essay, I lay out a completely different way of thinking about the evidence on industrial policy. I agree that there are many market failures in the world, that there can be important externalities from having a thriving industrial sector, and that potentially industrial policy can be a powerful tool to promote rapid economic growth and development. I also believe that this was the case in Korea, Taiwan,
China, and many of the other cases studied by the revisionists in the late 1980s. The balance of evidence suggests that these scholars were right to attribute a powerful causal role to industrial policy (though admittedly we have no definitive econometric evidence on this). However, I also believe that industrial policy can fail, as it did in Ghana in the 1960s and throughout Latin America from the 1940s onward. But the difference between these cases is not that the Japanese or the Koreans got lucky, were clever, or had better economists advising them; it was because the political equilibrium of these societies differed.

I argue that to begin to think properly about industrial policy, we need to start with a positive theory of such policy. It is remarkable the extent to which the economics literature on industrial policy has avoided proposing a positive theory of industrial policy as an equilibrium outcome. Take the reassessment of Soviet industrialization developed by Allen (2003). Allen argues that Stalin’s industrialization policy was optimal in a poor country with bad initial institutions. There is no argument, however, that would lead us to expect that the Stalinist regime would adopt an efficient policy (indeed, a vast mass of evidence suggests that efficiency was not high on the criteria that Stalin’s regime maximized; see Gregory and Harrison 2005). The normative approach to explaining policy assumes that politicians choose policy in a socially optimal way, something hard to believe about the Soviet Union in the 1920s and 1930s. Another striking example of this approach comes from the experience of policy reform in transition economies in the 1990s. While Russian reform was done badly, Chinese dual-track reform is characterized as clever (see Lau, Qian, and Roland 2000 for the argument that Chinese reforms were cleverer than Russia reforms). Of course, the Great Leap Forward during the 1950s is not viewed as clever, but as a big mistake. Although Shleifer and Treisman (2000) try to analyze Russian policy reforms as an equilibrium outcome, they still argue that they were (constrained) efficient. In fact, the evidence suggests that the reason that the Chinese did things differently than the Soviets was not because they were cleverer (see Shirking 1993; Naughton 2007). Chinese policy reform was an equilibrium outcome, as was Soviet reform. The fact that the transition policy differed was because, although both regimes may have been avowedly communist and (less avowedly) authoritarian, in fact the political equilibria in these societies were quite different. For example, the strength of the central state and the party was very different. The problem with industrial policy in Ghana in the 1960s was not that it was not clever or appropriate or needed, but rather that it was adopted in political circumstances where it had no chance of succeeding in actually industrializing the country.

To study any policy as an equilibrium outcome, to develop a positive theory of industrial policy, we need to understand the political economy. Variation in the adoption of industrial policy or in its success or failure has less to do with ideas or economists, although these can be important in particular circumstances, and much more to do with the nature of the political equilibrium in society—which interests are mobilized, what their interests are, what the political institutions are, and so forth. Understanding comparative policy is thus an exercise in comparative politics. This is a difficult exercise, and we are far from having a satisfactory framework that
can explain policy and outcome variation, but we certainly will never have such a framework until we start thinking about the issues in the right way.

Political economy focuses on developing a positive explanation for industrial policy. Does this leave any room for “industrial policy” in a more traditional sense? My argument should make clear that I think the answer to this is no. The problem of underdevelopment cannot be solved by economists coming up with better policies for poor countries to adopt or endlessly hoping for benevolent “leadership.” Economists have been proposing good policies for decades; the problem is that they are not adopted (just as poor countries do not adopt many of the other things that make countries rich, such as advanced technology). From my perspective, promoting industrialization—having an industrial policy—is an endogenous outcome of the political choices of a society. If interests and institutions are not aligned, then industry will not get promoted, whatever the normative consequences are. For outside economists or international institutions to stimulate industry, they have to take this into account.

This means two things. First, it means trying to change the nature of the political equilibrium in a direction more conducive to industrialization, for example, by strengthening the political influence of those groups who would benefit from this. This may seem like a radical idea, but the World Bank has few reservations in adopting policies aimed at strengthening the power of poor people (for example, with an eye to generating more accountability and better service delivery). Second, it means holding the political equilibrium constant, trying to find a way of crafting a pro-industry policy that will be incentive compatible for those holding power.

The paper proceeds as follows. First I discuss the normative approach to industrial policy and examine what the evidence suggests about its effectiveness. I then emphasize that the differences between the successes and failures are mostly due to differences in political economy. A section then sketches what a positive political economy theory of industrial policy should look like, followed by an examination of what such a positive theory would imply for how we should rethink our understanding of industrial policy. A final section concludes.

**Normative Industrial Policy**

The welfare economics argument in favor of industrial policy is straightforward and well understood. It is quite likely that market imperfections, externalities, and increasing returns exist, which suggests that various forms of industrial policy could be welfare improving and even necessary to create a modern manufacturing sector. The optimal form of the policy will depend on the nature of the market imperfection. For example, if industry is too small relative to the social optimum because of imperfections in the capital market, this does not provide a normative motivation for infant industry protection. Rather, policy should be focused on removing the specific market failure (though, of course, one has to bear in mind the Theorem of the Second Best, which suggests that in an economy with multiple market failures, removing one of them may make welfare worse rather than
better; Lancaster and Lipsey 1956–57). The simplest form of industrial policy—direct subsidies of industrial activity—would follow from welfare economics if industry generated positive externalities. In this case, a subsidy would be the canonical Pigouvian intervention. The choice of tariffs as the optimal intervention could arise in open economy models where industry again generates positive externalities and where international prices are such as to keep the industrial sector too small from a social point of view (as in the model of Matsuyama 1992). Nevertheless, a large literature has shown that even if one believes market failures are important, the actual policies chosen are rarely the ones that normative economic theory would predict—for example, inefficient instruments are used when efficient ones are available (see Coate and Morris 1995; Acemoglu and Robinson 2001). In this essay I leave this issue aside and simply observe that economic theory does provide a solid basis for believing that at least some forms of industrial policy can play a role in improving welfare and economic growth. In the next section I argue that there is evidence that this is indeed the case in practice. It should be obvious, however, that this type of theory does not provide a normative basis for many of the industrial policies we see, such as that adopted by the Soviet Union after 1928, although this is possibly the case under some very specific assumptions about the policy instruments available to the government.

The Experience of Industrial Policy

This section briefly discusses examples of failed and successful industrial policy. The main argument is not that industrial policy is always bad or always good. I believe, and I think the evidence suggests, that industrial policy has great potential to promote economic development. However, this potential can only be realized if the political environment is right. The examples are supposed to illustrate this perspective, which then is developed more systematically.

The poster children for those who advocate industrial policy are the East Asian “miracle” economies. A large literature documents this. Wade (1990, chs. 4–6) documents in great detail the case of Taiwan, China. There are many fascinating stories about how the government systematically intervened in the economy from the 1950s onward to promote industry. A famous one is how the government’s chief economic planner K.Y. Lin decided on the basis of a report by a U.S. consultant, J. G. White Engineering Corporation, that plastics was a suitable industry to develop (Wade 1990, 80). He then identified Y. C. Wang, a local businessman, as someone with the resources to do this, apparently through bank records. He then told Wang to start the business! The first factory was built under government supervision and given to Wang in 1957. Wang, subsequently head of the Formosa Plastics Group, went on to become a leading entrepreneur in Taiwan, China. Wade provides many examples of how the government intervened to stimulate both the quantity and quality of industry—for example, publicly destroying 20,000 light bulbs in Taipei to discourage poor-quality production (Wade 1990, 81). Although the private sector was developed in Taiwan,
China, the public sector was heavily involved in this industrialization drive, with as much as 60 percent of research and development expenditure attributed to the public sector in the 1980s (Wade 1990, 99) and a key role being played by the Industrial Policy Research Institute established in 1973. The institute played an important role in reducing technological dependence on the United States and launched “national strategic programs” in eight fields. The public sector introduced an export processing zone in 1965 and used many complementary instruments, such as credit, to stimulate exports (Wade 1990, 139–48). The government also started a large-scale integrated steel mill, as in Korea.

Figure 1, from Wade (1990, 111), sums up the periods in different industries where Wade judges that the state played a crucial role in leading particular industries. Some of these initiatives were not hugely successful. For example, unlike Japan or Korea, Taiwan, China, has not been able to develop an internationally competitive motor vehicle industry, despite a sustained attempt to do so. In other areas, such as semiconductors, these interventions have to be judged as very successful.

The evidence on the successful promotion of industry by the government in Taiwan, China, is impressive and convincing. This greatly bolsters the case for industrial policy. Unfortunately, however, there are many cases of unsuccessful industrial policy. Some of the best documented cases come from Sub-Saharan Africa. After independence many African countries adopted types of industrial policy, with some such as Ghana and Zambia announcing five-year plans and very ambitious targets. As in Taiwan, China, these programs were often led by the public sector. Unfortunately, in no Sub-Saharan African country did they generate an internationally competitive industry. Typically, while there was rapid capital accumulation, the industry that was developed was incredibly inefficient so that total factor productivity performance was abysmal.

One of the most detailed studies of the failure of industrial policy in Africa is Tony Killick’s seminal book about development in Ghana (Killick 1978). This should be required reading for anyone advocating industrial policy as a current solution to Africa’s problems. He discusses in great detail examples of industrial projects from the early 1960s. He shows in case after case that cost-benefit calculations were ignored and inefficient investment projects were undertaken. One example was a cattle-based industrial complex (Killick 1978, 231):

The footwear factory . . . would have linked the meat factory in the North through transportation of the hides to the South (for a distance of over 500 miles) to a tannery (now abandoned); the leather was to have been backhauled to the footwear factory in Kumasi, in the centre of the country and about 200 miles north of the tannery. Since the major footwear market is in the Accra metropolitan area, the shoes would then have to be transported an additional 200 miles back to the South.

Killick (1978, 231) remarks that this was an enterprise “whose viability was undermined by poor siting.”

Another startling example is the construction of a fruit canning factory “for the production of mango products, for which there was recognized to be no local market
FIGURE 1. State Leadership Episodes in Industries in Taiwan, China, 1950–90


Note: The lines refer to big leadership. The dates should be taken as rough approximations. The diagram does not show all industries that have experienced big leadership, nor is it necessarily complete for the industries shown.
[and] which was said to exceed by some multiple the total world trade in such items” (Killick 1978, 229). The government’s own report on this factory is worth quoting at some length (Killick 1978, 233):

Project. A factory is to be erected at Wenchi, Brong Ahafo, to produce 7,000 tons of mangoes and 5,300 tons of tomatoes per annum. If average yields of crops in that area will be 5 tons per acre per annum for mangoes and 5 tons per acre for tomatoes, there should be 1,400 acres of mangoes and 1,060 acres of tomatoes in the field to supply the factory.

The Problem. The present supply of mangoes in the area is from a few trees scattered in the bush and tomatoes are not grown on commercial scale, and so the production of these crops will have to start from scratch. Mangoes take 5–7 years from planting to start fruiting. How to obtain sufficient planting materials and to organize production of raw materials quickly become the major problems of this project.

Killick’s acerbic comment is that “it is difficult to imagine a more damning commentary on the efficiency of project planning,” stated a whole year before the factory was constructed. The problem under Nkrumah was not underinvestment in industry. Indeed, the consensus view is that the capital stock increased 80 percent between 1960 and 1965 (Killick 1978, 69), 60 percent of which was by the public sector (80 percent of nonresidential investment; Killick 1978, 170). The problem was in the way this investment was allocated.

It would be easy to present many pages of similar examples from Sub-Saharan Africa. Although the situation is less bad in Latin America, which has even had a few successes, particularly in Brazil, it is generally true that government stimulation of industry in Latin America has not been a success. Most government protection and subsidization of industry did not create internationally competitive firms; rather, it led to uncompetitive monopolies or oligopolies producing poor-quality goods for protected domestic markets.

The Political Economy of Success and Failure

What is the difference between the examples of the successful and unsuccessful industrial policy? Why did it work so well in Taiwan, China, but not in Ghana? Although there were certainly differences in the way that the policies were implemented (for example, inward looking versus outward looking), the main difference between these cases is political. After all, the Calico Acts seem to have been an example of successful import substitution, and one can argue that early import substitution in Taiwan, China, in the 1950s proved an important period that gave some subsequently dynamic firms a chance to get off the ground. So the strategy attempted, at least on paper, in Latin America and maybe even in Ghana could have worked. The fact that it did not was due to the politics of these countries.
To get a sense of my argument, let us return to Africa. Tangri (1999, 30) argues that the Industrial Development Corporation (INDECO) of Zambia failed to promote industrialization because

INDECO was subject to a series of ad hoc political directives on specific operational issues, including type and location of investments. Projects were undertaken on political considerations, although, as in the case of Mansa batteries, the feasibility study concluded that the project based in Mansa would be uneconomic. Moreover, projects such as the Chinese maize mill at Chingola were started without any feasibility study being undertaken; the decision was a purely political one, which led to the already planned and evaluated maize mill in Kitwe being abandoned. Directives were also issued regarding the location of projects. The locations of the Livingstone Motor Assemblers, Kapiri Glass Products, and Mansa Batteries, all subsidiaries of INDECO, were decided on the basis of providing employment outside the main urban areas. These and similar projects ran into difficulties for various reasons, partly because, being located in up-country centers, they were situated a long way from the main markets. Multi-million-dollar brick factories were set up under official directive in the rural areas at Kalalushi and Nega Nega, but transporting the bricks long distances to the construction sites raised their costs to uneconomic levels, with the result that the construction industry switched to the use of concrete blocks. Because of the declining demand for its products, the brick works at Nega Nega was forced to close down in 1979 and the factory at Kalalushi incurred large losses.

Tangri’s discussion of Zambian industrial policy suggests that the difference between INDECO and, say, Japan’s Ministry of International Trade and Industry was that in Zambia the choice of policy, project, and location was driven by political criteria, not economic ones.

In Ghana, as in Zambia, the motivation behind decisions to misallocate resources was clearly political. Rimmer (1969, 195) argues, “Projects were begun without feasibility studies and without competitive tendering. New enterprises were distributed among party functionaries as private fiefs, enabling them to give patronage to relatives, friends, and supporters.” Omaboe (1966, 460–61) concludes, “In Ghana the politicians are always ahead of the civil servants and planners in the general consideration and implementation of economic and social projects.”

In contradistinction, all of the literature on the East Asian miracle emphasizes how the economic bureaucracy was allowed to develop rational economic policies without having to adopt politically motivated projects or instruments.

**Positive Industrial Policy**

So industrial policy can work, but it may not work. The difference lies in the objectives and functioning of the institutions implementing the policies, and these are determined by the political system. This seems to be a much more important source of variation in success and failure than other factors, such as differences in binding constraints or in ideas about what to do. To explain the successes and failures of
industrial policy, we need to dig into these political factors. What was it about the politics of Taiwan, China, and the politics of Ghana that led to these outcomes? Why did their political equilibrium differ so much in ways that allowed one but not the other to implement a successful industrial policy? Important studies have addressed these questions, including Wade (1990), Haggard (1990), and Evans (1995).

To set the scene, I briefly discuss a theory of endogenous policy to give some background to my way of thinking about these issues (for an extensive discussion, see Acemoglu, Johnson, and Robinson 2005a). Economic policies (and more generally institutions that are usually also outcomes of collective choices) matter for economic growth because they shape the incentives of key economic actors in society; in particular, they influence investments in physical and human capital and technology and the organization of production. Economic policies determine not only the aggregate economic growth potential of the economy, but also the distribution of resources in the future. In other words, they influence not only the size of aggregate income, but also how income is divided among different groups and individuals in society.

Economic policies are endogenous and are determined as collective choices of the society. Clearly, there is no guarantee that all individuals and groups will prefer the same set of policies because, as noted, different policies lead to different income distributions. Consequently, there will be a conflict of interest over the choice of economic institutions. In such a situation it will be the distribution of political power in society that determines which institutions are chosen. The group with more political power will tend to secure the set of economic institutions that it prefers.

The distribution of political power in society is also endogenous, however. Following Acemoglu and Robinson (2006), I distinguish between two components of political power: de jure (institutional) and de facto. Here de jure political power refers to all types of power that originates from the political institutions in society. Political institutions determine the constraints on and the incentives of key actors in the political sphere. Examples of political institutions include the form of government—for example, democracy, dictatorship, or autocracy—and the extent of constraints on politicians and political elites. There is more to political power than political institutions, however. A group of individuals, even if they are not allocated power by political institutions (for example, as specified in the constitution) may nonetheless possess political power. Namely, they can revolt, use arms, hire mercenaries, co-opt the military, or use economically costly but largely peaceful protests in order to impose their wishes on society. I refer to this type of political power as de facto political power, which itself has two sources. First, it depends on the ability of the group in question to solve its collective action problem, that is, to ensure that people act together, even when any individual may have an incentive to free ride. Second, the de facto power of a group depends on its economic resources, which determine both its ability to use (or misuse) existing political institutions and also its option to hire and use force against different groups.
The composition of de facto and de jure power in society is what determines the actual power of a group or set of interests, and this will determine which economic policies arise. This vision emphasizes that those with power today take decisions not just to maximize their income today, but also to maintain their grip on power. These goals are often in contradiction. This can be for the simple reason that economic policies that increase the incomes of elites today may increase the incomes of opponents even more, thus influencing the future distribution of de facto power. It may also be that, as pointed out in the seminal study by Bates (1981), good economic policies are not good politics. In particular, although providing public goods may increase the incomes of the elite, staying in power may be better achieved by using redistributive instruments that can be targeted at supporters and withheld from opponents.

If the promotion of industry is the endogenous outcome of a political equilibrium, in what circumstances would the political equilibrium be likely to be propitious to it happening? If promoting industry was economically efficient in that it could create a Pareto improvement, then there is a sense in which any type of political system or any type of political leader might have an incentive to undertake it. Wittman (1989), for example, argues that in democracies political competition always leads to efficient policies. Nevertheless, Wittman’s arguments are subject to forceful counterarguments (Acemoglu 2003), and they do not appear to be relevant to the mostly nondemocratic regimes I have been discussing so far. Whatever led to successful industrial policy in Taiwan, China, it was not domestic political competition, since the government was run by the Kuomintang more or less as a one-party state. Of course, if industrialization produced economic success and wealth, then even authoritarian regimes ought to be in favor of it. The problem with this argument is that industrialization will almost certainly undermine the political status quo, as it did in both Korea and Taiwan, China.

These arguments suggest that industrialization is unlikely to be promoted by those who stand to benefit indirectly from it. Nevertheless, the extent to which this is true will certainly depend on circumstances. For example, both Korea and Taiwan, China, because of their colonial experiences under Japanese rule, had extensive histories of industrialization. When their governments began to promote industry in the 1960s, there were entrepreneurs and experience. Both Singapore and Hong Kong, China, had long histories of mercantile capitalism. In addition, in Korea and Taiwan, China, an extensive bureaucratic tradition played a key role in adopting and monitoring the policies. The politicians had to allow them to do this, but the fact that this capacity existed clearly could have influenced the probability of success of the policy and thus the incentive of the politicians to adopt it. Finally, these countries also had a lot of human capital. These circumstances suggest that the policy of promoting industry had a good chance of succeeding, and other things being equal, this would encourage any government to adopt it.

In addition to the factors that helped to make industrial policy potentially very effective, other factors reduced the political instability that might have flowed from it. Japan and Korea had long histories as independent consolidated polities, and
although in Taiwan, China, there was conflict between indigenous Taiwanese and mainlanders who came after the 1948 Chinese Revolution, there was a great deal of cultural and linguistic homogeneity. This was also true of Hong Kong, China, and Singapore. Moreover, all these societies were very egalitarian, with low levels of inequality, although this was partially the outcome of agrarian reforms. This low inequality along with the consolidated nature of the state implied that it might be feasible for incumbent political elites to promote rapid industrialization without being challenged for power (Acemoglu and Robinson 2006).

If indirect benefits are unlikely to generate industry promotion, one source of successful industrial policy will be the rise to political power of groups with specific investments in industry—those who stand to benefit directly from industrialization. This is exactly what Pincus (2009) argues about the Whig coalition in England. Many leading Whig politicians owned factories and nascent industrial enterprises, and they were in favor of policies that increased the value of their assets.

This argument is fine as far as it goes, but it misses a lot. This is because, as Adam Smith pointed out, one of the most attractive things for businessmen is a monopoly. The key characteristic of the Glorious Revolution is that it took place in the shadow of the 1623 Statute of Monopolies, which made it impossible to establish domestic monopolies in Britain. Thus the woolen and linen industries protected by the Calico Acts could not block the emergence of a domestic cotton industry that would compete with them. Moreover, the victorious Whig politicians could not themselves set up domestic monopolies after 1688.

These ideas are well illustrated by the experience of industrialization in Argentina. It is important to recognize that the policy switch in the direction of promoting industry in Argentina and Latin America more generally was an endogenous response to the formation of a new political coalition and was designed both to sustain it as well as to redistribute income to it. During the 1930s in Latin America, there was a huge change in the terms of trade, which led to a reallocation of resources into the urban and industrial sectors, where relative prices had risen. This shock, which came following movements toward greater democratic participation in many countries, greatly empowered urban and industrial interests and led to the emergence of a new politics. Along with this came the initiation of the ISI model of development.

This policy response took some time to emerge, and it did so in the context of the rise of new political movements and parties rooted in the changing economic and political landscape. In Brazil this was marked by the rise of Getúlio Vargas and the creation of the Estado Novo in 1937, the emergence of the Alianza Popular Revolucionaria Americana under Víctor Haya de la Torre in Peru, the presidency of Lázaro Cárdenas in Mexico between 1934 and 1940, and the rising power of Juan Perón in Argentina after 1943.

Although these new political forces promoted industry, the right way to think about this policy was that it directly benefited those newly empowered. Gerchunoff (1989) sums up Peronist economic policy in Argentina in the following way: “There was no specific and unified Peronist economic policy, much less a long-term development strategy. In spite of official rhetoric about a plan, the objective—and
at times exclusive—priority was . . . an economic order capable of maintaining the new distributive model.” There was an industrialization policy, because this was in the interests of those now choosing the economic policies, but the form that this policy took, while not as bad as that of Ghana in the 1960s, was not of a socially efficient form and certainly did not create the incentives necessary to stimulate rapid industrialization. In line with the evaluation of Gerchunoff (1989), Díaz Alejandro (1970, 126) concludes, “Peronist policies present a picture of a government interested not so much in industrialization as in a nationalistic and populist policy of increasing the real consumption, employment, and economic security of the masses—and of the new entrepreneurs. It chose these goals even at the expense of capital formation and of the economy’s capacity to transform.” Elsewhere, in a very relevant passage, Díaz Alejandro notes (1970, 65),

The main problem arises in that policies which are best from the viewpoint of economic efficiency (e.g. free, or nearly free, trade) generate an income distribution favorable to the owners of the relatively most abundant factor of production (e.g. land) and therefore strengthen the position of the traditional elite . . . long-run efficiency and a popular income distribution could only be reconciled by a sophisticated fiscal system, not an easy thing to achieve.

Compared to the Glorious Revolution, what was missing from the Peronist Revolution was the Statute of Monopolies. This discussion also indicates another large difference in the circumstances between the East Asian miracle economies and Latin America or Africa. In none of the East Asian countries were there strong rural interests who could either oppose industrialization or be exploited by urban groups. Singapore and Hong Kong, China, had no hinterland, and in Japan, Taiwan (China), and Korea agrarian reforms removed the power of rural elites.

Why did Britain have a Statute of Monopolies, but not Argentina? These differing outcomes have deep roots in the historical development of the institutions of these societies. The policy outcomes chosen in Latin America in the 1930s and 1940s were the path-dependent outcome of a process of institution formation that went back to the colonial period (Engerman and Sokoloff 1997; Acemoglu, Johnson, and Robinson 2001). Latin American colonial societies developed economic institutions aimed at creating and extracting rents from indigenous peoples and slaves. This greatly benefited a narrow elite at the expense of society and was responsible for the huge levels of inequality in that region.

The political coalition that formed in Latin America in the 1930s and the policies it adopted were highly path dependent, reflecting the institutional history of Latin America, with a focus on rent creation and extraction. This led to monopolies and a highly clientelistic political strategy, which mirrored that of the previous oligarchies. The path of institution creation in Britain was very different, and in the late medieval and early modern period a series of shocks and structural changes radically changed the balance of power in society toward one that not only empowered those with different policy interests, but also created a series of de facto and de jure checks and balances out of which measures such as the Statute of Monopolies arose (see Acemoglu, Johnson, and Robinson 2005b; Pincus 2009).
The successful promotion of industry in East Asia is a consequence of the historical development path there. In addition to some of the structural factors that influenced the costs and benefits to political elites of industrial policy, one more factor merits emphasis. Korea and Taiwan, China, both former Japanese colonies, were heavily influenced by the experience of defense modernization on which Japan had embarked after the Meiji Restoration in an attempt to avoid being colonized by Europeans. Such a goal also heavily motivated Sun Yat-sen and the Chinese nationalist movement, which took over Taiwan, China, in 1948. Moreover, their precarious international position and intense competition with the Democratic People’s Republic of Korea and China created large incentives for the governments to promote industry. These factors made it far more imperative for Korea and Taiwan, China, to develop a strong industrial sector.

A New Approach to Industrial Policy

My interpretation of this evidence suggests that failed industrial policy, like bad policy more generally, is the consequence of perverse political incentives. The successful promotion of industry therefore requires changes in the political equilibrium in such a way as to align the incentives of the politically powerful with those of society. This may be achieved by changes in political institutions, or it may be achieved by endogenous changes in the balance of de facto power in society. The successful industrial policy of East Asian countries reflects the very different political equilibrium that emerged historically in this part of the world compared to Latin America or Sub-Saharan Africa.

The historical roots of these development paths and the endogenous nature of industrial policy do not lend themselves to simple policy proposals because it is not clear how to replicate the political equilibrium of Korea or Taiwan, China, elsewhere in the world. My discussion should have made clear that many unique factors are associated with the experience of these economies. Industrial policy—promoting industry—requires an understanding of the political equilibrium of a society, the actors and their interests, the political institutions, de facto and de jure power, and how these pieces all fit together. In such a situation there will be many potential ways to promote industry. One would be intervening to strengthen those with a vested interest in this outcome. Another would be to change political institutions to try to induce greater political competition in the hope that this would stimulate more socially efficient policies. An alternative to nudging the political equilibrium would be to try to work within it, for example, by creating incentives for existing political elites to change policy. Without sketching in any detail what a framework for this type of industrial policy would look like, I hope I have made clear why we need to think in this way.

Conclusions

In this essay I have tried to make a simple argument. There are good reasons to believe from economic theory that industrial policy may be socially desirable and
may stimulate economic growth and development. Some salient examples support this idea. Some industrial policies seem to have worked. Yet some have not. I have argued that the difference between these cases lies in the politics of the different societies that adopted the policies. Like any socially desirable policy, industrial policy may not necessarily be an outcome of a political equilibrium, and if it is, it may not be of a desirable form (it may be more about redistributing income or political power than promoting economic growth). Thus the reason that industrial policy failed in so many African countries in the 1960s is the same as the reason that economic policies were generally very bad in that region: policies were driven by the desire to maintain political power, and this was generally inconsistent with economic growth. The difference between these cases and less extreme ones in Latin America and East Asia is that the political equilibrium in East Asian countries was very different. The geopolitical and historical situation in East Asia was distinct. Finally, I have argued that, if this political economy perspective is correct, then this entails thinking about industrial policy in a new way. It is not sufficient just to propose good economic policies; one must propose a way in which they will be endogenously chosen by those with the political power to do so.

References


Comment on “Industrial Policy and Development: A Political Economy Perspective,” by James A. Robinson

BERT HOFMAN

Professor Robinson has written an eloquent, insightful, relevant, and somewhat provocative paper. The paper draws on his rich historical and institutional knowledge and is in many ways an application of the analysis of his seminal paper “The Colonial Origins of Economic Development” (Acemoglu, Johnson, and Robinson 2001) and his more recent paper “Persistence of Power, Elites, and Institutions” (Acemoglu and Robinson 2008). The paper’s core thesis is that the outcome of industrial policies in the end depends on the political equilibrium in which these policies are applied and that this political equilibrium explains why those policies work well in some countries, including some East Asian countries, and not in others, including some African and Latin American countries. The implication for policy advisers such as myself, the paper argues, is that analyzing and understanding the political economy of a country are as important as, if not more important than, the technical-economic design of industrial policies themselves.

The core insight of the paper is a major step in the direction of a positive theory of industrial policy, one that is very welcome. Rather than arguing that (a) government failures, erroneous economic theories, or rent seeking undermine industrial policies or, in contrast, that (b) industrial policies can work in the presence of a number of technical success factors such as openness and a competent bureaucracy, the paper proposes that the underlying power structure is the true determinant of the outcome of industrial policies. This power structure, Robinson argues, is the outcome of long historical processes, often dating back to colonial times. It can change, of course, through shocks such as wars, revolutions, and technology, and the paper presents examples of such shocks, which can in the end lead to a new equilibrium that is conducive to good outcomes of industrial policies. But dramatic shocks and revolutions...
Aside, the policy adviser is usually confronted with an environment in which power elites are well established and the outcome of any policy is predetermined to be in the interests of those power elites. These interests often diverge from the “common good” of the greatest advantage for the largest number of people. While entirely plausible, the core thesis is hard to prove or disprove empirically. Professor Robinson does provide several examples of industrial policies gone right and wrong. But because the “interests of the power structure” are not directly measurable, the thesis becomes almost tautological: “The outcome of industrial policies was the way it was because the underlying power structure determined it to be the way it has turned out to be.”

Fortunately, there are also examples of industrial policies that are clearly in the interests of the power elite and at the same time also provide benefits to the general public. The dual-track reforms in China may well be an example of that. Reforming the Chinese economic system at the margin—leaving the planned system in place, while allowing the market to work with above-plan output—was not only an astute political compromise between reformers and old-fashioned communists, but also appropriate economic policies for the broader masses. For that matter, even the “crony capitalism” under Suharto’s Indonesia, which clearly benefited a select group of entrepreneurs in a disproportionate way, also lifted more than half of the population out of poverty in the course of 30 years. So Robinson’s conclusion that all successful industrial policies serve the power elite need not be as dismal as it first sounds. Indeed, good outcomes, he argues, can still be achieved in an environment of selfish power elites, but again, this makes Robinson’s theory rather Panglossian: “Every outcome of industrial policy is the best possible outcome imaginable given the underlying political equilibrium.”

Robinson raises some important questions on the role of the policy adviser, including the role of the World Bank. He recommends that the economic policy adviser should pay attention to political economy factors and the specific political equilibrium that will determine the outcome of the policy actions he recommends. Being a policy adviser myself, I am slightly bemused by the recommendation because this is indeed what we do on a daily basis, even though considerations of political economy may not always be prominent in official reports for obvious reasons. The fact of the matter is, however, that such analysis is hard to do, and the consequences of policy recommendations—in other words, the trajectory from one equilibrium to the next—are rarely obvious, even if one tries to analyze the actions from a political economy perspective.

And indeed, the equilibrium is not static. Take China again: even if more standard market-based reforms would have been economically beneficial in 1978, they were simply not politically feasible, and the World Bank, which started to engage with China in 1980, was well aware of this. But by the early 1990s, after 15 years of reforms and the creation from scratch of something that might be called an entrepreneurial class, much more was possible, and the Bank (and others) felt much more confident in recommending the type of market-based reforms that were approved in 1993. Or take Indonesia, where industrial policies that by and large drew on the comparative advantage of the country worked well in the first two decades of Suharto’s
reign, but became increasingly captured by connected insiders and in the end contributed to the severe economic crisis (and the fall of the Suharto government). Hard as it is to determine the political equilibrium so as to offer the right policy advice, it becomes even harder as this political equilibrium is dynamic and sometimes changes under the influence of the very economic reforms recommended in the past.

Nevertheless, I believe that more insights into the political economy and dynamics of the countries we work in are indeed most welcome, and, as mentioned, Professor Robinson’s paper makes a valuable contribution toward those better insights. In the meantime, though, we have to do with our amateurish insights and, indeed, trial and error, because the political power structure often only reveals itself in the outcome of policies advised, and the results, as Professor Robinson and others show elsewhere, are often unintended or unexpected.

Finally, Professor Robinson seems to put little value on ideas and economic theory as determinants of the outcomes of industrial policy. In fact, he hardly bothers to define those policies, saying “all policies that influence industry.” This is somewhat surprising, as those ideas have evolved dramatically over the last five decades. In the 1950s it was a respected view that the state should take “control over the commanding heights of the economy,” and import substitution—so maligned afterward—was considered a viable development strategy. By the 1980s, under the influence of new academic ideas as well as disappointing outcomes of the earlier strategies, and I may add Professor Anne Krueger’s seminal contributions (Krueger 1974) and intellectual legacy at the World Bank, structural adjustment and “getting the prices right” were the new paradigm, preached and practiced in many diverse political environments. Of course, Robinson’s point is that this was wrong, as both strands of policy recommendations neglected the political equilibrium. Maybe so, but the technical design of the policies themselves still has a role to play. Even with a conducive political environment, a policy adviser can advise the wrong thing and get the recommendations accepted. Indeed, even the success stories in East Asia show numerous examples of wrong turns and dead ends in the pursuit of industrial development. The difference with the unsuccessful examples is perhaps that they did not persist in doing the wrong thing and instead dropped unpromising ideas and switched to better ones. Thus my conclusion is that there is more than political economy to the art of industrial policy making.

References


Industrial Policy: Can We Go Beyond an Unproductive Confrontation?

HA-JOON CHANG

This paper attempts to go beyond what the author sees as an unproductive confrontation between the proponents and the opponents of industrial policy and to take the debate on industrial policy forward. After discussing some issues related to conceptualizing and assessing industrial policy, the paper discusses most (although not all) of the key issues emerging from the industrial policy debate. They include the wisdom or otherwise of targeting, the feasibility of the state “beating the market,” political economy questions, bureaucratic capabilities, performance measurement (especially export targets), the importance of export-related industrial policy, and the implications of changing global policy environment.

Few topics in development economics, and indeed in economics as a whole, have caused more heated controversy than industrial policy. Not just its effectiveness and generalizability, but also its definition and very existence have been debated. Its opponents have declared its nonexistence, irrelevance, ineffectiveness, and demise many times, but it refuses to go away. For this to be the case, there has to be something more than the irrepressible human tendency to search for a magic solution for their problems. This paper aims to go beyond what I see as an unproductive confrontation between the proponents and the opponents of industrial policy and to explore how we can take the debate forward. I cannot claim to be impartial in this endeavor, as I have been a party to this debate. I will, however, do my best to find common ground and extract some theoretical and policy lessons from both sides of the debate.

Ha-Joon Chang is reader in the political economy of development on the Faculty of Economics at the University of Cambridge.

Annual World Bank Conference on Development Economics 2010, Global
© 2011 The International Bank for Reconstruction and Development / The World Bank
The Industrial Policy Debate: Conceptual Issues and Neglected Facts

Before discussing what I think are the main lessons from the industrial policy debate, I briefly review the debate itself. While I cannot avoid pronouncing judgments on at least some of the arguments advanced during the debate, the main purpose of the review is not to declare scores. It is to highlight some conceptual issues and neglected facts that help us to see the debate from what I hope to be a broader, but more pragmatic, point of view.

Literally interpreted, industrial policy should mean policy that affects industry, in the same way in which agricultural policy means policy that affects agriculture and monetary policy means policy that affects monetary variables. And indeed, many commentators on industrial policy on both sides of the argument follow this definition (see Chang 1994, 58–61, for some examples).

However, when we talk about “industrial policy,” the majority of us do not mean any policy that affects industry, but a very particular type of policy that affects industries. This type of policy is commonly known as selective industrial policy or targeting—namely, a policy that deliberately favors particular industries over others, against market signals, usually (but not necessarily) to enhance efficiency and promote productivity growth.

Industrial policy in this sense is usually associated with the development experiences of Japan and other East Asian economies (the Republic of Korea, Taiwan, China, and Singapore) in the post–World War II period. As I explain below, however, industrial policy, even in this narrow sense, has been practiced well beyond such time and place. Even so, let me start with the debate on post–World War II East Asian industrial policy, as this is what has framed our current thinking on industrial policy.

The modern debate on industrial policy started in the late 1970s, with the rise of Japan. Although the practice of (selective) industrial policy had been noticed among the scholars of postwar French economic policy in the 1960s, it was seen as a part of the broader exercise of “indicative planning” (Shonfield 1965; Cohen 1977). With the debate on Japan, industrial policy was brought to center stage, not least because Japan was the first country that used the term “industrial policy” (sangyo seisaku) to mean selective industrial policy. By the late 1980s, it came to be widely accepted that strong industrial policy was also practiced in Korea; Taiwan, China; and (in a very different way) Singapore; which had until then been thought to be free trade, free market economies.

In the early days of the debate on industrial policy in East Asia, some denied its very existence. Some of this denial was out of sheer unwillingness to recognize any fact that goes against one’s deep-held beliefs. For example, free trade economist Bela Balassa argued, as late as 1988, that the role of the state in Korea “apart from the promotion of shipbuilding and steel . . . has been to create a modern infrastructure, to provide a stable incentive system, and to ensure that government bureaucracy will help rather than hinder exports” (Balassa 1988, S286). However, more often it was based on an honest misunderstanding of the ways in which industrial policy worked in these countries. For example, Trezise (1983) argued that Japan did not have much industrial policy on the “objective” ground that its industrial subsidies
and government loans as a proportion of GDP were below the Organisation for Economic Co-operation and Development (OECD) average.

However, subsequent debate revealed that industrial policy in East Asia involved a lot more than handing out subsidies and providing trade protectionism (tariffs, import bans, quotas, and domestic regulations at least partially intended to curb imports). Industrial policy measures in East Asia included (a) coordination of complementary investments (the so-called Big Push); (b) coordination of competing investments through the regulation of entry, “investment cartels,” and (in declining industries) negotiated capacity cuts; (c) policies to ensure scale economies (for example, licensing conditional on scale of production, emphasis on the infant industries starting to export early on, and state-mediated mergers and acquisitions); (d) regulation of technology imports (for example, screening for overly obsolete technologies and caps on technology licensing royalties); (e) regulation of foreign direct investment (for example, entry and ownership restrictions, local contents requirements, technology transfer requirements, and export requirements); (f) mandatory worker training for firms above a certain size, in order to resolve the problem of underinvestment in the training of skilled workers due to the possibility of poaching; (g) the state’s role in providing venture capital and incubating high-tech firms; (h) export promotion (export subsidies, export loan guarantees, and marketing help from the state trading agency); and (i) government allocation of foreign exchanges, with top priority going to capital goods imports (especially for export industries) and the bottom priority going to luxury consumption goods imports.

The debate on the existence and the definition of industrial policy in East Asia alone has suggested two important points that are important to bear in mind when assessing industrial policy in general.

First, the extent of industrial policy cannot be identified purely in terms of quantifiable measures, especially those that involve financial transfers. As can be seen from the above list, many industrial policy measures do not involve any financial transfer, possibly except in the most roundabout general-equilibrium sense. By looking only at quantifiable indicators, we significantly underestimate the extent and the depth of industrial policy, both at the sectoral level and at the economy-wide level.

Second, we cannot assess the impact of a country’s industrial policy solely on the basis of the performance (however measured) of the “targeted” sectors (World Bank 1993 and Lee 1996 are the two most frequently cited examples along this line). Looking at sectors separately leaves us to ignore the impacts of “super-sectoral” industrial policy measures that address issues such as complementarities, linkages, and externalities among sectors.1

Of course, as the critics of industrial policy rightly point out, the mere coexistence of industrial policy, however widespread, and rapid industrial economic development in East Asia does not prove that the former has caused the latter. As they point out, it is possible that these countries could have grown even faster, had they not used industrial policy (Pack and Saggi 2006).

This is logically possible, but if that were the case, these countries must have had some country-specific “countervailing forces” that were so powerful that they canceled out all the harmful effects of market-distorting industrial policy and still
generated the highest growth rates in human history (6–7 percent annual growth rate in per capita income over four decades). I find this highly implausible. Are these skeptics seriously suggesting that, without industrial policy, these powerful countervailing forces would have made the East Asian countries grow at 9, 10, or even 12 percent, when no country in history has ever grown faster than 7 percent for an extended period, with industrial policy or not?

No convincing story as to what these countervailing forces are has been offered. Culture (leading to a high savings rate, strict work ethic, and high-quality bureaucracy), the legacy of Japanese colonialism (leading to exceptionally high literacy and a broad industrial base), and cold war politics (leading to exceptionally high foreign aid and special access to the U.S. market) are frequently cited, but none of them even passes the minimum factual tests (Chang 2007, ch. 9, on culture; Chang 2006 on Japanese colonialism and the cold war).2

Of course, as Pack and Saggi (2006, 268) point out, it is impossible to prove that East Asia could have done better or worse without industrial policy, as “the relevant counterfactuals are not available.” However, not all counterfactuals are equally plausible, and the counterfactual supposed by the critics of industrial policy is highly implausible. This nudges us toward the conclusion that industrial policy worked in East Asia.

Moreover, contrary to what many of its critics believe, industrial policy success was not confined to late-twentieth-century East Asia. There is quite a lot of evidence outside late-twentieth-century East Asia that further strengthens (but once again cannot “prove”) the case for industrial policy. There are three such sets of evidence.

First of all, if we broaden our spatial horizon, we realize that successful industrial policy experiences in the late twentieth century are not confined to East Asia. I have already mentioned the French industrial policy, but quite a few other European economies, such as Finland, Norway, and Austria, also pursued (selective) industrial policy, often with even greater success than France, during this period (Katzenstein 1985). Certain local governments in Italy (Emilia-Romagna) and Germany (Baden-Württemberg) also pursued effective industrial policy, promoting particular “industrial districts” through directed credits (from local banks, often owned by the local government), support for research and development (R&D), and export marketing help (Piore and Sabel 1984). All these countries had high growth rates between the 1950s and the 1980s, although this is not to say that industrial policy was solely responsible for their growth.3

While championing the free market ideology during this period (although not before that), the U.S. government also ran a huge (if somewhat wasteful) industrial policy program under the guise of R&D support for defense and public health. Between the 1950s and the 1980s, the U.S. federal government financed anywhere between 47 and 65 percent of national R&D spending, compared with around 20 percent in Japan and Korea and less than 40 percent in several European countries, such as Belgium, Finland, Germany, and Sweden (Mowery and Rosenberg 1993, 41, table 2.3 for the United States; OECD data for the other countries).4 Many of the industries where the United States still has a technological edge would
not have developed, or even emerged at all, without public funding of R&D. These include aircraft, computers, microchips, the Internet, and genetic engineering.

Second, going back in time, there are even more stories of industrial policy success. Contrary to the popular myth, in the nineteenth and the early twentieth centuries, all of today’s rich countries, except for the Netherlands and (before World War I) Switzerland, practiced significant degrees of protectionism for substantial periods of time (see table 1; see Bairoch 1993 and Chang 2002 for further details). Although these tariffs were not as systematically calibrated as those used in the late twentieth century in East Asia (and other countries), they were definitely parts of (selective) industrial policy insofar as they were deliberately different across sectors. In addition

<table>
<thead>
<tr>
<th>Country</th>
<th>1820b</th>
<th>1875b</th>
<th>1913</th>
<th>1925</th>
<th>1931</th>
<th>1950</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>R</td>
<td>15–20</td>
<td>18</td>
<td>16</td>
<td>24</td>
<td>18</td>
</tr>
<tr>
<td>Belgium</td>
<td>6–8</td>
<td>9–10</td>
<td>9</td>
<td>15</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>Canada</td>
<td>5</td>
<td>15</td>
<td>—</td>
<td>23</td>
<td>28</td>
<td>17</td>
</tr>
<tr>
<td>Denmark</td>
<td>25–35</td>
<td>15–20</td>
<td>14</td>
<td>10</td>
<td>—</td>
<td>3</td>
</tr>
<tr>
<td>France</td>
<td>R (20)</td>
<td>12–15</td>
<td>20</td>
<td>21</td>
<td>30</td>
<td>18</td>
</tr>
<tr>
<td>Germany</td>
<td>8–12</td>
<td>4–6</td>
<td>13</td>
<td>20</td>
<td>21</td>
<td>26</td>
</tr>
<tr>
<td>Italy</td>
<td>—</td>
<td>8–10</td>
<td>18</td>
<td>22</td>
<td>46</td>
<td>25</td>
</tr>
<tr>
<td>Japan</td>
<td>R</td>
<td>5</td>
<td>30</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Netherlands</td>
<td>6–8</td>
<td>3–5</td>
<td>4</td>
<td>6</td>
<td>—</td>
<td>11</td>
</tr>
<tr>
<td>Russia</td>
<td>R</td>
<td>15–20</td>
<td>84</td>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Spain</td>
<td>R</td>
<td>15–20</td>
<td>41</td>
<td>41</td>
<td>63</td>
<td>—</td>
</tr>
<tr>
<td>Sweden</td>
<td>R</td>
<td>3–5</td>
<td>20</td>
<td>16</td>
<td>21</td>
<td>9</td>
</tr>
<tr>
<td>Switzerland</td>
<td>8–12</td>
<td>4–6</td>
<td>9</td>
<td>14</td>
<td>19</td>
<td>—</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>45–55</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>—</td>
<td>23</td>
</tr>
<tr>
<td>United States</td>
<td>35–45</td>
<td>40–50</td>
<td>44</td>
<td>37</td>
<td>48</td>
<td>14</td>
</tr>
</tbody>
</table>

Source: Chang 2002, 17, table 2.1, largely based on Bairoch 1993, 40, table 3.3, except for Canada, which is from Taylor 1948, 102–08, 398.

Notes: R = numerous and important restrictions existed, making average tariff rates not meaningful. — = not available.

a. World Bank (1991, 97, box table 5.2) provides a similar table, partly drawing on Bairoch 1993. However, the World Bank figures, although in most cases very similar to Bairoch’s, are unweighted averages, which are obviously less preferable to the weighted average figures that Bairoch provides.

b. These are very approximate rates and give a range of average rates, not extremes.

c. Austria-Hungary before 1925.

d. In 1820, Belgium was united with the Netherlands.

e. According to the estimate by Nye (1991), the average tariff rate, measured by customs revenue as a percentage of net import values, in France during 1821–25 was 20.3 percent, as against 53.1 percent for Britain, which is in line with the 45–55 percent range estimated by Bairoch.

f. The 1820 figure is for Prussia only.

g. Before 1911, Japan was obliged to keep tariff rates low (up to 5 percent) through a series of unequal treaties with the European countries and the United States. World Bank (1991, 97, box table 5.2) gives Japan’s unweighted average tariff rate for all goods (not just manufactured goods) for 1925, 1930, and 1950 as 13, 19, and 4 percent, respectively.
to tariff protection, many of these countries provided subsidies to promote targeted industries, set up state-owned enterprises or public-private joint ventures for risky projects, regulated foreign direct investments, and implemented many other measures of industrial policy during this period (Chang 2002, 2007).

Britain and the United States—the supposed homes of free trade—had the world’s highest levels of tariff protection (45–55 percent) during their respective catch-up periods—from the mid-eighteenth to the mid-nineteenth century for Britain and from the mid-nineteenth century to the mid-twentieth century for the United States.

This was no coincidence. Robert Walpole, the so-called first British prime minister, is credited to have been the first person to launch a comprehensive infant industry program in 1721 (Brisco 1907). Walpole strongly influenced Alexander Hamilton, the first Treasury secretary of the United States, who first developed the theory of infant industry protection (Hamilton 1791). The targeted protections that Germany and Sweden provided to their nascent heavy industries in the late nineteenth and early twentieth centuries are well known, but even Belgium, one of the less protected economies, provided targeted protection. In the mid-nineteenth century, when Belgium’s average industrial tariff was around 10 percent, the textile industries had tariffs rates of 30–60 percent and the iron industry had rates of 85 percent (Milward and Saul 1977, 174). At least for the 1870–1913 period, there is even evidence of a positive correlation between tariff rate and rate of growth (O’Rourke 2000; Vamvakidis 2002; Clemens and Williamson 2004).

Third, the long-term historical experiences of the developing countries also provide some food for thought regarding industrial policy. Drawing on numerous studies that show a positive cross-section correlation between a country’s degree of “openness” (variously measured) and its growth performance, the mainstream consensus is that industrial policy in developing countries since the 1960s has not worked. Even if we ignore many criticisms of these cross-section econometric studies (Rodriguez and Rodrik 2000; Chang 2005) and accept such a conclusion, the time-series evidence tells us a rather different story.

Until the 1870s, most of today’s developing countries practiced free trade, either because they were colonies or because they were bound by the so-called unequal treaties that deprived them of tariff autonomy and imposed a low, uniform rate of tariff (3–5 percent). However, their growth performance during this period was very poor (see table 2). When the Latin American countries gained tariff autonomy in the 1870s and the 1880s, their per capita income growth rate shot up from 0.1 percent during 1820–70 to 1.8 percent during 1870–1913, making it one of the two fastest-growing regions in the world during the latter period (table 2).

The growth performance of the developing countries during the “bad old days” of import substitution industrialization was a vast improvement over their performance before and, more important, has not been matched since the 1980s, when they abandoned much of their industrial policy. Per capita income in developing countries grew at 3 percent a year during 1960–80 (World Bank 1980, 99, table SA.1). Their growth rate fell to just above half that (1.7 percent) in the next 20 years (calculated from World Bank 2002), when these countries liberalized and
opened up their economies. The slowdown in growth was particularly striking in Latin America and Sub-Saharan Africa, two regions that most faithfully implemented market-oriented reforms during this period. Per capita income in the two regions grew at 3.1 and 1.6 percent a year, respectively, during 1960–80 (World Bank 1980, 99, table SA.1), compared with 0.5 and –0.3 percent, respectively, during 1980–2004 (calculated from the World Bank and the United Nations Development Programme data sets).

Individually, this evidence as well as the evidence about the East Asian experience discussed earlier do not prove anything. However, taken together, they raise some difficult questions for the skeptics of industrial policy. If industrial policy was not confined to East Asia in the late twentieth century, it becomes difficult to downplay its role in East Asia by resorting to some region- and time-specific “countervailing forces.” Even if many countries that have used industrial policy did not succeed, the fact that few of today's rich countries have become rich without industrial policy makes us wonder whether a good industrial policy may be a necessary, although not sufficient, condition for economic development. Looking at all these sets of facts together, we have to wonder, if industrial policy is so bad, how is it that in every era, the fastest-growing economies happen to be those with a strong industrial policy: Britain during the mid-eighteenth century and mid-nineteenth century; the United States, Germany, and Sweden during the late nineteenth and the early twentieth century; East Asia, France, Finland, Norway, and Austria in the late twentieth century; and China today.

Although the weight of evidence is, on the whole, on the side of (intelligently conducted) industrial policy, we do not need some absolute “proof” of its merit, either way, in order to take the debate forward. As far as we can agree that the chance of success for industrial policy is more than negligible, we can still have a productive debate on how to make it work better, even if we cannot agree on the

<table>
<thead>
<tr>
<th>Region</th>
<th>1820–70</th>
<th>1870–1913</th>
<th>1913–50</th>
<th>1950–73</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western Europe</td>
<td>0.95</td>
<td>1.32</td>
<td>0.76</td>
<td>4.08</td>
</tr>
<tr>
<td>Western offshoots a</td>
<td>1.42</td>
<td>1.81</td>
<td>1.55</td>
<td>2.44</td>
</tr>
<tr>
<td>Japan</td>
<td>0.19</td>
<td>1.48</td>
<td>0.89</td>
<td>8.05</td>
</tr>
<tr>
<td>Asia excluding Japan</td>
<td>–0.11</td>
<td>0.38</td>
<td>–0.02</td>
<td>2.92</td>
</tr>
<tr>
<td>Latin America</td>
<td>0.10</td>
<td>1.81</td>
<td>1.42</td>
<td>2.52</td>
</tr>
<tr>
<td>Eastern Europe and the former Soviet Union</td>
<td>0.64</td>
<td>1.15</td>
<td>1.50</td>
<td>3.49</td>
</tr>
<tr>
<td>Africa</td>
<td>0.12</td>
<td>0.64</td>
<td>1.02</td>
<td>2.07</td>
</tr>
<tr>
<td>World</td>
<td>0.53</td>
<td>1.30</td>
<td>0.91</td>
<td>2.93</td>
</tr>
</tbody>
</table>

Source: Maddison 2001, 126, table 3-1a.

a. Australia, Canada, New Zealand, and the United States.
exact “batting average” of industrial policy. Therefore, in the rest of the paper, I discuss some of the lessons that we have learned (or should have learned) from the actual experiences of and the theoretical debates on industrial policy and suggest some ways forward, both theoretically and in terms of pragmatic policy.

**What Have We Learned? Lessons from the Experiences and the Debates**

In this section, drawing on the industrial policy debate and adding some of my own take on it, I explore how we can make industrial policy work better. I look at issues surrounding (a) targeting; (b) whether the state can “beat the market”; (c) political economy; (d) bureaucratic capabilities; (e) performance measurement; (f) export; and (g) changing global environment. Although it is quite wide ranging, this list leaves out some key issues in the industrial policy debate, especially the challenges of building productive capability and the problems due to adjustment costs (on these issues, see Lin and Chang 2009; Dosi, Cimoli, and Stiglitz 2009).

**The Question of Targeting: Selective vs. General Industrial Policy**

After at least three decades of intense debate on industrial policy, few people would deny that there are instances where state intervention in industrial development is justified. However, many would argue that industrial policy should be “general” (or “functional”) rather than “selective” (or “sectoral”). They argue that the state should concentrate on providing things like education, R&D, and infrastructure that benefit all industries equally but are likely to be underprovided by the market, rather than trying to “pick winners” by favoring particular sectors or even firms. In other words, they reject industrial policy in the usual sense, while not rejecting the idea that the state can (and should) overcome market failures in relation to industrial development.

The first problem with this view is that the distinction between selective and general industrial policies cannot be taken very far. In a world with scarce resources, every policy choice you make, however “general” the policy may look, has discriminatory effects that amount to targeting. This point is easier to see in relation to R&D—a government giving out R&D subsidies implicitly favors the more R&D-intensive high-tech sectors—but it also applies to infrastructure and education, at least to the higher ends of them. We do not build some abstract infrastructure; instead we build either a road between the horticultural export region and an airport or a railway between a steel town and a seaport. Building the railway, instead of the road, means that the government at least implicitly favors the steel industry. Likewise, we do not educate some generic engineers; instead we educate either chemical engineers or electronics engineers. Therefore, a government that provides more funding to electronics engineering departments than to chemical engineering departments is implicitly favoring the electronics industry. The only policies that may be called truly “general” are policies regarding basic education and health;
calling them industrial policies stretches the concept beyond reason. Thus seen, selectivity and targeting are involved in virtually every (broadly defined) industrial policy measure. The only real difference is that of degree.

If targeting is unavoidable, can we at least say that the less targeted a policy is, the better it is? We cannot. The more targeted a policy is, the easier the monitoring of the beneficiaries is, and therefore the “leakages” are going to be less. Indeed, mainstream economists recommend more precise targeting in social policy for this reason (on targeting in social policy, see Mkandawire 2005). Why is this point not considered in relation to industrial policy? Of course, targeting has its costs. For example, too precise a targeting may, in a world with fundamental uncertainty, be bad because it “puts all eggs in one basket.” Or it may make lobbying easier. Or it may make the beneficiaries too easy to identify, making it difficult for the government to maintain the necessary myth that its policies are impartial. And so on.

The debate on industrial targeting needs to move to a higher level. While accepting its potential problems, the inevitability of targeting should be acknowledged. We should drop the pretense that we can “not target” and instead try to attain the best possible degree of targeting, which may differ across industries and countries. We should stop thinking that there is a linear relationship, positive or negative, between the degree of targeting and policy success: some degree of targeting is inevitable, and, while more targeting may be desirable, too much may not be good. Perhaps we should think in terms of “targeting within universalism,” as in the debate on social policy (Skocpol 1991, as cited in Mkandawire 2005, 23), rather than “targeting vs. universalism.”

Can the State “Beat the Market”? Ability, Information, and Perspective

One of the classic arguments against infant industry protection (and by extension any selective industrial policy) is that the private sector would have promoted the industry in question if it were genuinely worth promoting (Baldwin 1969). Given that the government officials by definition know less about business than do businessmen, the argument goes, it is inevitable that their decisions are likely to be of lower quality than those made by businessmen. In other words, the state cannot “beat the market.”

However, there are quite a few examples in history where government officials made investment decisions that blatantly went against market signals, sometimes even using state-owned enterprises as vehicles, only to build some of the most successful businesses in history. The four decades of protection, subsidies, and ban on foreign direct investment (FDI) in the Japanese automobile industry before its world market success, the entry of Korea into the steel industry through a state-owned enterprise (Pohang Iron and Steel Company [POSCO]) in 1968 (when the country’s per capita income was only 5 percent that of the United States), or Brazil’s entry into the aircraft industry, once again through a state-owned enterprise (Embraer) in 1969 (when its per capita income was only 8 percent that of the United States) are only some of the most spectacular examples (Chang 2002, ch. 2; 2007, ch. 5; 2008).
These cases are euphemistically known as (the government officials successfully correcting) “capital market failures,” but it would be far more honest if we admitted that the state can sometimes beat the market. Against this, Pack and Saggi (2006) admit that there are capital market failures but argue that the solution should be found in developing the banking sector, “perhaps by allowing foreign financial intermediaries into the country” (Pack and Saggi 2006, 270) that have “modern techniques of evaluating individual projects and managing the riskiness of their overall portfolio” (Pack and Saggi 2006, 285), rather than in industrial policy. However, this suggestion rings hollow today, when those “modern techniques” have created perhaps the biggest financial mess in human history.

More important, we do not need the assumption that government officials are omniscient or even that they are cleverer than capitalists in order to advocate industrial policy. The point is that many (although not all) of the “superior” decisions made by the state were made not because the government officials were omniscient or cleverer than businessmen but because they could look at things from a national and long-term point of view, rather than a sectional, short-term point of view.

It is because they saw things from a national point of view that the East Asian government officials could prevent domestic firms from outbidding each other in licensing foreign technologies or could take externalities into account and encourage things like exporting and training beyond what seems “rational” to individual businessmen. It was because they could take a more long-term view that the Korean and the Brazilian states could set up firms like POSCO or Embraer, ventures that “rational” private sector firms did not want to touch with a barge pole.

If we do not need to assume that bureaucrats are omniscient in order to justify industrial policy, we can have a much more meaningful discussion on how to improve the quality of industrial policy. If some bureaucrats are indeed better businessmen than capitalists, we can learn how to run better industrial policy by asking what kind of people they are, how they make the decisions, and how they implement them—in the same spirit with which we read books by and about famous businessmen. If bureaucrats make better decisions simply because they have more “systemic” perspectives, we can perhaps improve private sector decisions by encouraging the formation of industry associations or a national business association. We should also discuss whether there are organizational forms that encourage even more long-term-oriented and more systemic thinking in the bureaucracy.

**Political Economy: Leadership, Bureaucracy, and Power**

The ability of the state to improve on market outcomes thanks to its more systemic and longer-term perspective is only the starting point of running a good industrial policy. The debate has revealed a key difference between stories of industrial policy success and failure: the former had states that could impose strict discipline on the recipients of its support (Toye 1987; Amsden 1989; Chang 1994; Evans 1995). Since the state conducting industrial policy is at least partially suspending market discipline, it has to supply the necessary discipline itself. If government supports are seen
as “handouts,” rather than as “advances” for the delivery of good performance in the future, the recipients of government support will have no incentive to perform.

Many complex political economy issues have been debated over the years. Here, I present what I consider to be the key lessons at three levels of political economy—that of political leadership, internal control structure of the state, and the power of the state vis-à-vis the rest of the society.

First, let us look at the nature of the top leadership. It cannot be assumed a priori that the leaders running a particular state are interested in economic development, whether through industrial policy or not. In order to appreciate this point, we do not need to go to the extreme and believe that all top political leaders are “predators” interested only in personal wealth and aggrandizement (although some may well be). Even if they are interested in economic development, the leaders may have a “wrong” vision. They may be looking backward rather than forward—Thomas Jefferson and his followers were vehemently opposed to Hamilton’s policy, as they wanted to preserve a society made up of respected landlords and yeoman farmers (plus the slaves). Or the political leaders may be hostile to private sector development, as many developing-country leaders were in the 1960s and the 1970s. Or, as many nineteenth-century liberal politicians did, they may think that doing nothing, other than protecting private property, is really the best policy.

Second, even if the political leaders are interested in promoting economic development through industrial policy, they need to impose that vision on the rest of the state apparatus. While in theory the state is a hierarchical organization, in practice the wish at the top does not always permeate through the hierarchy. Once again, we do not need to go to the extreme and assume that government officials are seeking only their own self-interests (for example, the self-seeking bureaucracy approach of Niskanen) to see this point. There will naturally be some degree of self-seeking, but many real-life bureaucrats are dedicated public servants. However, there would still be problems arising from clashing visions (for example, the bureaucrats may be more conservative than the political leaders), turf wars between different groups within the bureaucracy, “tunnel vision” (which specialized organizations are wont to develop), internal coordination failures (coming from poor organizational design or the emergence of new issues that cut across the existing organizational structure), and many other reasons.

Third, even if everyone in the government, from the top leader down to the lowliest clerk, shares the same vision about industrial policy, the state still has to be able to impose its will on other agents in the society. Needless to say, the feasibility of this differs across countries (and across issues, even within the same country). In some extreme cases, the state may not even have full control of its claimed territories. In some developing countries, the state may not be able to implement policies because it is short of manpower and resources, especially when it tries to influence an industry with numerous small firms, where monitoring is costly. Even when the state has enough enforcement capabilities, some private sector agents will attempt to neutralize or even pervert policies through lobbying and bribing. Some groups may have such influence in the society that the state does things that they want or refrains from doing things that they do not want, even without explicit lobbying or
outright corruption, as we are witnessing these days in relation to the financial industry in countries like the United States or the United Kingdom.

Thus seen, between accepting the need for industrial policy and implementing it there is a huge range of political economy problems. Indeed, when considering that many developing countries are run by flawed leaders presiding over a politically weak and internally fragmented state, it seems difficult to imagine how industrial policy, even if it were “correct,” can be implemented well in a developing country.

However, we should not let the best be the enemy of the good. The existence of numerous political economy problems should not make us believe that therefore we have to wait for a perfect state to emerge before doing anything.

In the real world, successful countries are the ones that have managed to find “good enough” solutions to their political economy problems and gone on to implement policies, rather than sitting around bemoaning the imperfect nature of their political system. Of course, in the long run, these countries’ successes also owe something to their investments in the improvement of their state, including the quality of the bureaucracy, interest group organizations, and the very nature of their society, but when they started their development, they started with highly deficient political systems (on the history of institutional development in today’s rich countries, see Chang 2002, ch. 3).

Quite a few of them, including some of the successful “industrial policy states,” overcame their political obstacles to effective statecraft in situations that did not instill much hope. For example, between the fall of Napoleon and the end of the Second World War, the French state was notoriously laissez-faire, ineffectual, and conservative. However, this changed after the war, with the rise of Gaullisme, the establishment of the planning commission, and the foundation of the École Nationale d’Administration, the famous school for elite bureaucrats (Cohen 1977; Kuisel 1981). For another example, the Kuomintang (Nationalist Party) bureaucracy was one of the most corrupt and inefficient in modern history when it ruled mainland China, but after being forced to migrate to Taiwan, China, by the communists, it was transformed into a highly efficient and relatively clean one. This was done through a gradual but deliberate process of building “islands of competence” and then giving them greater responsibilities as they succeeded and increased their legitimacy and status within the bureaucracy, finally replacing much of the old bureaucracy with the new one (Wade 1990).

So, rather than assuming away the political economy problems (as some proponents of industrial policy have done) or using them as an excuse for policy inaction (as some opponents of industrial policy have done), we should find ways to devise imperfect but workable solutions to those problems. In order to take the debate forward, we have to improve our understanding of issues such as (a) how to form and deploy effective political visions to inspire individuals and groups to act in a concerted manner; (b) how to build nations and communities out of disparate groups that may even have a very long history of mutual hostility and mistrust; (c) how to work out social pacts and build lasting coalitions behind them; (d) how to accept but improve the customs and organizational routines in the bureaucracy; and (e) how to minimize socially harmful lobbying and bribing, while maximizing the
flows of information between the states and the private sector. In order to address these issues, we economists need to go beyond the usual boundaries and work with practitioners (politicians, government officials, and businessmen) as well as academics from other fields (political science, sociology, anthropology, psychology, and cultural studies).

**Bureaucratic Capabilities: Important, but Not in the Way We Think**

However willing and strong the state may be and however “correct” its vision may be, policies are likely to fail if the government officials implementing them are not capable. Difficult decisions have to be made with limited information and fundamental uncertainty, often under political pressure from inside and outside the country, which requires decision makers with intelligence and adequate knowledge.

On this ground, people have argued that “difficult” policies like (selective) industrial policy should not be tried by countries with limited bureaucratic capabilities. And it is for this reason that the World Bank (1993) recommended the Southeast Asian countries (Thailand, Malaysia, and Indonesia) as models of industrial policy for other developing countries. In these countries, industrial policy was quite circumscribed partly in recognition of the relatively low quality of their bureaucracies.

At the general level, I cannot dispute the proposition that capable bureaucrats are needed to design and implement good policies. I agree too that a policy that has succeeded spectacularly in one country can turn into a mess in another country in the hands of incompetent bureaucrats, in the same way in which the same recipe can result in a masterpiece by a top chef, a pleasing dish by a good cook, and a culinary disaster by an incompetent cook. I also agree that policies differ in their difficulties and therefore need to be chosen according to the relevant government’s capabilities.

Unfortunately, these sensible points are often assembled into the policy world equivalent of the “do not try this at home” warning that accompanies the demonstration of difficult and dangerous tricks in television shows. It is argued that industrial policy is so difficult that it should never be tried by countries that do not have an East Asian-style high-quality bureaucracy, which in effect means all developing countries. Is this acceptable?

First, one critical assumption behind the “do not try this at home” argument is that industrial policy is exceptionally difficult. However, the assumption is made without any theoretical reasoning or empirical evidence. For example, World Bank (1993) assumes that policies getting the “fundamentals” right, such as human capital, agriculture, and macroeconomic stability, are easier than industrial policies, but there can be no such presumption. First, different governments have competences in different areas: the Japanese government was good at industrial policy, but not at macroeconomic policy in the 1990s. Second, the ease of a policy will also depend partly on its scale. For example, promoting a few industries may be a lot easier than organizing a mass education program. Third, it will also depend on the number of agents involved in the policy. Trying to coordinate investments among
a few large firms may be a lot easier than organizing a countrywide distribution of subsidized fertilizer that involves millions of small farmers who are not organized into cooperatives and are scattered all over the country.

Second, another implicit assumption behind the “do not try this at home” argument is that industrial policy requires sophisticated knowledge of economics, as exemplified by the following comment by Alan Winters, one-time head of the Research Department at the World Bank and now the chief economist of the U.K. government’s Department for International Development: “The application of second-best economics needs first-best economists, not its usual complement of third- and fourth-raters” (Winters 2003, as cited in Stiglitz and Charlton 2005, 37). But is this true? The interesting thing is that while the East Asian bureaucracies were staffed by smart people, they were certainly not “first-best economists.” Indeed, most of them were not even economists. The Japanese economic officials that engineered the country’s “miracle” were mostly lawyers by training. Until the 1980s, what little economics they knew were mostly of the “wrong” kind—the economics of Karl Marx and Friedrich List—rather than neoclassical economics. In Taiwan, China, most key economic bureaucrats were engineers and scientists, as is the case in China today. Korea also had a high proportion of lawyers in its economic bureaucracy until the 1970s, while the brains behind the Korean Heavy and Chemical Industrialization (HCI) program in the 1970s, Oh Won-Chul, was an engineer by training. Both Korea and Taiwan, China, had rather strong, albeit officially unacknowledged, communist influence in their economic thinking until the 1970s.

Third, many people who advance the “do not try this at home” argument believe that high-quality bureaucracies are very difficult to build and that the East Asian countries were exceptionally lucky to have inherited them from history. However, a high-quality bureaucracy can be built pretty quickly, as shown by the examples of Korea and Taiwan, China, themselves. Contrary to the popular myth, these two economies did not start their economic “miracles” with high-quality bureaucracies. For example, until the late 1960s, Korea used to send its bureaucrats for extra training to, of all places, Pakistan and the Philippines. Taiwan, China, had a similar problem of generally low bureaucratic capabilities in the 1950s and most of the 1960s. These countries could construct a high-quality bureaucracy only because they invested in training, organizational reform, and improved incentive systems. In addition, there was a lot of “learning-by-doing.” By trying out industrial policy early on, the East Asian bureaucrats could more quickly pick up and improve the capabilities they needed to run industrial policy effectively. In other words, there has to be some “trying at home,” if you aspire to become good enough to appear on television with your own act.

Last, but not least, the fact that something is “difficult” cannot be a reason not to recommend it. When it comes to personal advancement, we go to the other extreme and encourage our youngsters to aspire to become the best of the best (by reading biographies and what not), when most of them are going to end up as production-line workers or shop assistants rather than prime ministers or business tycoons. And when it comes to institutions, even developing countries
are routinely told to adopt “best-practice” or “global-standard” institutions used by the richest countries, when many of them clearly do not have the capabilities to run effectively the American patent law system, the British accounting system, or the Scandinavian welfare system. But when it comes to industrial policy, countries are told to aim low and not to try at all or, at best, to try to learn from the Southeast Asian countries rather than the East Asian countries (or other rich countries, for that matter). I am all for people warning against the risks involved in “aiming too high,” but why should countries aim low only when it comes to industrial policy?

The critics of industrial policy have made an important contribution by highlighting the importance of bureaucratic capabilities in implementing industrial policy. However, this does not mean that a country with a low-quality bureaucracy should not aspire to implement “difficult” policies, like industrial policy (if it is difficult). Capabilities can be increased over time through deliberate investment and through learning-by-doing (of the “difficult” policies). To be more productive, therefore, the debate should focus on issues such as the following: (a) Exactly why is, or isn’t, industrial policy more difficult than other policies? (b) If it is more difficult than other policies, can it be made “easier” by learning from “best practices”? (c) If the capabilities needed are not those in mainstream economics, what bureaucratic capabilities are needed for good industrial policy? (d) How can we build those capabilities most quickly and cheaply?

**Performance Measurement: Difficulties and Mitigations**

Even with a willing, strong, and capable state, imposing discipline on the recipients of state support is not a straightforward business. At the most general level, we can say that the recipients should be rewarded for good performance and punished for bad performance, but translating that principle into practice is not easy, not least because of the difficulties involved in measuring performance.

Especially when industrial policy is comprehensive, as it was in the case of East Asia between the 1950s and the 1980s, objective performances become difficult to measure, as virtually all prices are “distorted.” There will also be efforts by the recipients of state help to manipulate the performance indicators. These are real and serious problems, but the industrial policy debate has revealed that there are ways to overcome them.

First of all, when launching an industrial policy program, performance targets should be clearly specified and the reporting requirements on them set so that the recipients cannot weasel their way out of bad performance. Publicly announcing the targets will make their manipulation more difficult, although it will reduce policy flexibility.

Second, the targets should be set in consultation with the business community, so that they are realistic and do not simply reflect some bureaucratic dreams. However, they should not be set purely on the basis of what entrepreneurs say, as they are likely to overstate the difficulties and understate the strengths, so there have to be independent third opinions provided by technical experts, academics,
journalists, and the like. The deliberation councils that were used in Japan and, less effectively, in Korea show how this process can be managed (Johnson 1982; Dore 1986; World Bank 1993).

Third, targets need to be revised along the way: they may prove to be too easy or too difficult, or they may be unexpectedly affected by external shocks. In particular, it is important for governments to acknowledge mistakes quickly and to change policies, as they did in East Asia. Having said that, government flexibility can be abused by lobbying, so too much flexibility should be avoided.

Fourth, in industries where export is possible, export performance should be given high status as a measure of performance, as in the East Asian countries, especially in Korea. Export performance indicators are far less open to manipulation by the recipients of state support than are indicators of domestic market performance, especially when the firms in question have significant market powers.

Fifth, policy makers need to pay more attention to the trends in performance indicators rather than to their absolute levels at any given point in time. This is particularly important in programs with a long time horizon, such as the plan to develop the automobile industry in Japan and Korea, which took literally two or three decades before bearing even the first fruits.

Once again, it is time for the debate to move on. Rather than debating whether setting and enforcing effective performance targets is possible (as it certainly is), we should concentrate on questions like the following: (a) What performance indicators should be used for which industries? (b) How do we set credible performance targets without becoming too rigid about them? (c) How does the government listen to the private sector without becoming beholden to it? (d) How do we operate with a long but not infinite time horizon?

**The Importance of Export-Related Industrial Policy**

I have just discussed the role of export in helping the state to better discipline the recipients of its support by providing a relatively objective and hard-to-manipulate indicator of performance, but export has other important roles to play in the conduct of industrial policy in developing countries.

To put it bluntly, economic development is impossible without good export performance. Economic development requires importation of advanced technologies, in the form of either machines or technology licensing, which need to be paid for with foreign currencies. Unless a country is so small or so strategically located that it receives disproportionately large amounts of foreign aid or foreign direct investment, it will simply have to export its way out of poverty.

The failure to promote export enough is one of the key reasons why the Latin American industrial policies were not as successful as those in East Asia. In the Latin American countries, economic growth kept hitting the balance of payments constraints. Even with its huge export machinery and massive government support for exports (subsidized bank loans, tariff rebates for imported inputs used for exports, export marketing support from the state trading agency), Korea found it impossible
to export enough to finance its rapid rate of capital accumulation until the late 1980s, running constant trade deficits.

So far, I am singing from the same hymn sheet as the World Bank. However, saying that export is the key to economic development is not to say that developing countries should liberalize their trade and closely follow their comparative advantages.

Of course, at the beginning of its economic development process, a country should try to increase its exports from its existing industries and other “nontraditional” industries where it has comparative advantages (for example, salmon in Chile, coffee in Vietnam, cut flowers in some African countries). The widespread view is that these industries do not need any export help because they are in line with the country’s comparative advantage, but this is wrong.

Export success requires significant industrial policy even for comparative advantage–conforming industries, especially if they are “nontraditional” industries, where new productive capabilities may have to be built. The basic problem is that export markets have high fixed costs of entry, which smaller firms and farmers, who are likely to dominate these comparative advantage–conforming industries in developing countries, may not be able to bear. Direct export subsidies can offset the costs of entry, but these are now banned by the World Trade Organization (WTO), except for the least developed countries, so the help should be provided through other channels.

First, state marketing help can be crucial, especially for smaller exporters, as exemplified by the Japan External Trade Organization and the Korea Trade Promotion Corporation, known as the Korea Trade–Investment Promotion Agency since 1995, as well as the Danish agricultural marketing boards in the early twentieth century (on the Danish case, see Chang 2009). Second, the state could share risk with exporters through schemes like loan guarantees for exporters and insurance for payment defaults. Third, it can help exporters, especially small producers, to meet the high quality standards required in the export markets. This can be done through, for example, controlling the quality of export products, providing advice on sanitary and phyto-sanitary requirements in the agricultural export markets, and providing subsidized extension services for small farmers and small manufacturing firms engaged in exporting. Fourth, the state can help indirectly by providing legal and financial support for cooperative arrangements among exporters for the joint provision of marketing services, R&D, storage facilities (warehouses and refrigerated stores), processing facilities (creameries and slaughterhouses), and transport facilities (refrigerated trucks) (see Chang 2009).

In the longer run, if it is to continue the momentum of its export success, a country will have to rely on more than its comparative advantage–conforming industries. Especially given the nature of the industries with which developing countries are likely to start their export drives, export growth is likely to peter out soon after the initial stage and even a small rise in wages (which the export success will bring) may undermine the country’s position in the world market. Sooner or later, it will have to upgrade its export industries into comparative advantage–defying industries, which requires even stronger industrial policy.
A good example of this is Korea. In the 1950s, Korea’s main exports were tungsten ore, fish, seaweed, and basic textiles and garments. In the 1960s, the government developed nontraditional labor-intensive export industries like wigs, plywood, shoes, and cheap electronics assembly, with the help of massive export support programs, while upgrading existing export industries, especially the textile and garment industries. By the early 1970s, however, many of these export industries, especially plywood and wigs, were hitting the wall, so the country launched the HCI Program, developing industries such as shipbuilding, steel, petrochemicals, automobiles, and high-end electronics as export industries, despite the fact that it did not have comparative advantage in those industries at the time. Without these industries, however, Korea would not have sustained its export growth momentum, and thus its ability to grow quickly, beyond the 1970s.

Indeed, what truly distinguishes the East Asian countries from other developing countries is not that they had “freer” trade than others. After all, they had plenty of protectionism. Average industrial tariff rates were 30–40 percent both in Korea and Taiwan, China, until the 1970s, while both of them had numerous nontariff trade barriers. The real difference is that in East Asia, free trade, export promotion (which is, of course, not free trade), and infant industry protection were organically integrated, both in cross-sectional terms (so some industries are always subject to each category of policy, sometimes more than one at the same time) and over time (so the same industry may be subject to more than one of the three over time).

Therefore, while emphasizing the importance of export for economic development, we need to abandon the “export promotion vs. import substitution” dichotomy that has dogged the industrial policy debate for far too long. We need to debate how exactly to mix free trade, export promotion, and infant industry protection—across sectors and over time—in a manner that helps a country to upgrade its industrial structure and grow quickly. We also need to discuss the factors that determine the optimal mix of these three types of trade policy and the timing of switching between them.

**A Changing Global Environment**

Considerable changes have happened to the global economy since the heyday of industrial policy between the 1950s and the 1980s. Two mutually reinforcing sets of changes have happened—changes in the global business environment and in the rules of global trade and investment—that people argue have made industrial policy almost impossible to implement. Later I look more closely at the changes in trade and investment rules, which are more directly relevant to our discussion, but first I make some brief comments on the changes in the business environment.

Changes in the global business environment can be broken into two related components. The first is the increase in the importance of foreign direct investment and the other is the so-called “global business revolution,” which has led to enormous concentration at the top tiers of the global value chain, which in turn has forced consolidation down the value chain.
The rising importance of FDI has made a lot of commentators think that it is very difficult, if not impossible, for countries to use “nationalistic” industrial policies for fear that transnational corporations will move away.

While the relative importance of FDI has increased enormously since the 1990s, the changes are not as dramatic as they are often thought to be. The pace of change may slow down, as the current crisis is forcing transnational corporations to retrench. In the longer run, the trend may even be reversed. After all, globalization was reversed once during the interwar period. Moreover, the mobility of transnational corporations differs enormously across industries and in relation to different countries, so the feasibility of nationalistic policies depends on the industry and the country (Chang 1998, 2007, ch. 4). A country with a large domestic market and good supply-side conditions (skills and infrastructure) may implement a very nationalistic policy in an industry with low mobility (for example, automobiles and steel), but a country without those conditions cannot do the same, especially when it comes to high-mobility industries (for example, garments and shoes). Also, empirical studies reveal that industrial policy, such as imposing a performance requirement on transnational corporations, is not as important in influencing FDI decisions as market conditions (the size and growth of the domestic market), infrastructure, or quality of the labor force (Chang 1998; Kumar 2005).

Since the 1980s but accelerating since the 1990s, there has been an enormous increase in industrial concentration, starting with the top of the global value chain and increasingly cascading down the chain, a process that has been named the “global business revolution” by Nolan, Zhang, and Liu (2008). This revolution has enormously raised the entry barriers to higher-end industries faced by developing countries.

However, industry concentrations go up and down in the long run, so it is not certain that the current trends will continue forever. Especially given the turmoil in the world economy today, new spaces higher up the value chain may open up for developing-country companies as existing producers decline (particularly visible at all levels of the automobile industry at the moment), while recent mergers and acquisitions, which have come to account for an increasingly higher share of FDI, may be undone. In the long run, when new industries emerge, opportunities arise for new entrants (for example, East Asia and the electronics revolution), so some developing countries may emerge as major players in some new industries in the future. Moreover, value chains have become more “chopped up” and internationalized, adding further new opportunities for developing-country producers. Developing countries can now enter lower segments of those industries that used to be vertically integrated and located only in the rich countries (for example, automobiles), although this route is highly unlikely to allow them to reach the top of the chain.

More directly relevant for this paper than the changes in business landscape are the changes in global rules of trade and investment. The use of many of the classic tools of industrial policy are now either banned or significantly circumscribed by the
WTO. Quantitative restrictions (for example, quotas) have been banned altogether. Tariffs have been reduced and “bound.” Export subsidies are banned, except for the least developed countries. Most other subsidies have become open to countervailing duties and other retaliatory measures. New issues, like regulations on FDI and intellectual property rights, have been brought under the jurisdiction of the WTO, making it difficult for countries to “borrow” foreign technologies for free or to put performance requirements on transnational corporations.

Thus seen, the WTO has certainly made industrial policy more difficult to implement. However, the constraints imposed by the WTO should not be exaggerated.

First, even on paper, the WTO by no means obliges countries to abolish all tariffs, and many developing countries have “bounded their tariffs” (or set tariff ceilings) at considerably high levels. Of course, if the rich countries have their way in the current nonagricultural market access negotiations of the Doha Round in the WTO, industrial tariffs in the developing countries are, at 5–10 percent, likely to fall to the lowest level since the days of colonialism and unequal treaties (Chang 2005, 4). However, this is yet to happen.

Second, there are still provisions for emergency tariff increases (“import surcharges”) on two grounds. The first is a sudden surge in sectoral imports, which some countries have already used. The second is the overall balance of payments problem, for which almost all developing countries qualify and which quite a few countries have also used. Since countries have discretion over the coverage and the levels of emergency tariffs that are meant to lessen the balance of payments problem, there is still room for targeting particular industries.

Third, not all subsidies are “illegal” for everyone. For example, the least developed countries are allowed to use export subsidies. Subsidies for agriculture, regional development, basic R&D, and environment-related technology upgrading were explicitly allowed (non-actionable in WTO parlance) until 1999. Even though the last three have become “actionable” since 2000, not a single case has been brought to the dispute settlement mechanism since then, suggesting that there is an implicit agreement that they are still acceptable. Moreover, the subsidy restrictions only cover “trade-related” ones, which means that “domestic” subsidies can be used (for example, subsidies for investments in equipment and in particular skills).

Fourth, the TRIPS (trade-related intellectual property rights) agreement has made technology absorption more expensive for developing countries (Chang 2001). However, this mainly affects the middle-income countries. The technologies that many developing countries need are often the ones that are too old to have patents.

Last, the TRIMS (trade-related investment measures) agreement bans measures like local contents requirements and trade-balancing requirements, which had been successfully used by both the developed and the developing countries (Kumar 2005). However, countries can still impose conditions regarding the hiring of local labor (a good way to create technological spillover effects), technology transfer, and the conduct of R&D in the host country. They can also provide targeted subsidies, directed credits, and tailor-made infrastructure (measures that Singapore and Ireland have used to attract FDI into “targeted” industries; Chang 2004), provided that these do not violate the most-favored-nation provision (Thrasher and Gallagher 2008).
Of course, even though the WTO rules allow quite a lot of industrial policy measures, especially for the least developed countries, this space is in practice highly constrained by other international factors. In the case of the least developed countries, the conditions attached to bilateral and multilateral aids and loans, on which they are quite dependent, significantly constrain their industrial policy space. Many developing countries are also parties to bilateral and regional trade and investment agreements, which tend to be even more restrictive than the WTO agreements (Thrasher and Gallagher 2008).

So, all in all, the range of industrial policy measures that developing countries can use has become considerably smaller, compared to the heyday of industrial policy in the 1960s and the 1970s, partly because of the changing global business landscape but more importantly because of the changes in global rules of trade and investment. However, there is still room for maneuver for countries that are clever and determined enough. Moreover, especially with the current world financial crisis, the global business landscape can change significantly, opening up unexpected possibilities of moving up and across global value chains for at least some developing countries. The global rules of trade and investment are not unalterable laws of nature. They can be, and should be, changed if they are found wanting.

Conclusions

In this article, I have tried to find some ways to overcome what I consider to be an unproductive confrontation between the proponents and the opponents of industrial policy and to take the debate forward by exploring some common ground between the two groups.

In the first substantive section, I briefly reviewed the debate on industrial policy, emphasizing that we need to look beyond East Asia between the 1950s and the 1980s in order to deepen our understanding of industrial policy. While I explained why I think that the weight of evidence is in favor of industrial policy, my aim was not to declare who has “won” the debate (which is impossible to prove anyway), but to establish the minimum common empirical understanding for a more productive debate. This common understanding is that industrial policy can work, sometimes spectacularly well, but it can also fail, sometimes miserably. This is a moderate proposition, which most (albeit not all, I realize) people on both sides of the debate can live with and on the basis of which they can engage in a pragmatic debate on how to make industrial policy work better.

In the second, and main, substantive part of the paper, I looked at most (although not all) of what I regard as the key issues emerging from the industrial policy debate. I discussed some theoretical questions regarding whether targeting is desirable and whether the state can “beat the market.” I looked at implementation issues, ranging from “big” political economy problems, through to questions surrounding bureaucratic capabilities, down to nitty-gritty issues related to the measurement of performance. After emphasizing the importance of export performance as an indicator of performance, I talked about the critical importance of export policy, which
requires not just free trade but a mixture of free trade, export promotion, and infant industry protection. I then discussed how the changing global business landscape and, more important, the recent changes in global rules of trade and investment are affecting the feasibility of industrial policy and how these will evolve themselves.

While I could not avoid making some partial statements—I did not want to hide my status as a well-known advocate of industrial policy—my main purpose in this article was to plea for “thinking outside the box” and finding the common grounds for people on both sides of the debate.

We, on both sides of the debate, have focused too much on “grand” things like the Big Push, when much of real-life industrial policy has been about “boring” things, like getting the scale of production right and providing export marketing services. This is not surprising because most practitioners of industrial policy over the last two centuries have been pragmatic people who did not know many fancy economic theories. Some theoretical issues that both the proponents and the opponents consider to be critical dissolve into thin air once seen from a pragmatic point of view (for example, targeting or bureaucrats as businessmen). Many proponents of industrial policy do not fully appreciate how critical export is for the success of industrial policy, while many opponents do not fully appreciate how export success also requires industrial policy. We often let sensible worries (such as political economy and bureaucratic capabilities) degenerate into a recommendation for inaction, letting the best become the enemy of the good. Real-life success stories were often based on “good enough” compromises rather than perfect solutions.

Once the adversaries in the debate abandon theoretical grandstanding and focus on more practical issues, there are vast and fertile middle grounds to explore. This is not to say that there will not be disagreements. However, at least the two sides can have productive debates on pragmatic issues without thinking about destroying each other. Would that be too much to ask?

Notes

1. In addition to being unable to address the super-sectoral dimensions of industrial policy, World Bank (1993) and Lee (1996) have the following problems. Looking at 38 industrial sectors (basically at the three-digit level) in Korea between 1962 and 1983, Lee (1996) found largely negative correlation between a sector’s receipt of government support (tariffs, nontariff barriers, tax incentives, and subsidized loans) and its performance, measured by various indicators (such as labor productivity, total factor productivity, or TFP, and capital intensity). The study should be commended for collecting much detailed data and looking at more than TFP, which has a lot of conceptual and practical problems; however, focusing on quantifiable measures, it could not capture many important aspects of industrial policy, even at the sectoral level (such as getting scale economies right and coordinating competing investments). Moreover, when infant industries require 10, 20, or even 30 years to mature, assessing Korean industrial policy in 1983 gives a bias against it—Korea’s main industrial policy drive, the Heavy and Chemical Industrialization Program, was launched only in 1973. Third, by stopping in 1983, the study underestimates the performance of the young heavy and chemical industries, which suffered disproportionately in the 1979–82 economic downturn prompted by exogenous factors.
(oil price rise and monetarist policies in the United States). World Bank (1993), looking at Japan, Korea, and Taiwan, China, assumed that sectors (defined at the two-digit industry classification level) with higher value added components or higher capital intensity were supported more by the government, thus obviating (perhaps unintentionally) the problem of relying only on quantifiable variables. It tried to correlate a sector's value added component and capital intensity with its performance (measured, unfortunately, only in terms of TFP) and found positive correlation only in Japan. However, the East Asian government targeted sectors at a much more disaggregated level than the two-digit one and never on simple grounds like capital intensity or value added component. For example, the textile industry in Korea, whose good performance the World Bank takes as a sign that “neglected” industries did quite well, was in fact one of the most “targeted” sectors until the mid-1980s because of its role as the main foreign exchange earner (Chang 1995, ch. 3, app.; also see Rodrik 1994).

2. Let me provide some basic factual refutation of these “countervailing forces” arguments, a full treatment that is beyond the scope of the paper. Before their economic development, the East Asians were typically described as lazy, unenterprising, individualistic people, “living for today” (see Chang 2007, ch. 9). Korea's savings rate on the eve of its economic miracle was barely 5 percent; it started rising after growth took off. At the end of the Japanese colonial rule, the literacy rate in Korea was only 22 percent and its industrial base was smaller than that of Ghana (Chang 2006). After the 1950s, Korea and Taiwan, China, did not receive an exceptionally high amount of foreign aid (Chang 2006). As far as I know, no one has provided any concrete evidence for the “special market access” story. Until the 1980s, Korea and Taiwan, China, were buying up textile quotas from other developing countries that could not even fill their Multi-Fibre Agreement quotas for the United States, showing that even if it existed, special market access could not provide big enough export markets to these two countries.

3. Of the 16 largest OECD economies studied by Maddison (1989) between 1950 and 1987, the seven fastest-growing economies, in per capita terms, were Japan (6 percent), Austria (3.9 percent), Germany (3.8 percent), Italy (3.7 percent), Finland (3.6 percent), Norway (3.4 percent), and France (3.2 percent).

4. The share of federal government in total R&D spending was 5.36 percent in 1953, 56.8 percent in 1955, 64.6 percent in 1960, 64.9 percent in 1965, 57.1 percent in 1970, 51.7 percent in 1975, 47.2 percent in 1980, 47.9 percent in 1985, and 47.3 percent in 1989 (estimated).

5. Irwin (2002) argues that this correlation was driven by high tariffs imposed for revenue reasons in the New World countries (the United States, Canada, and Argentina in his sample) that were growing quickly for other reasons (such as rich natural resource endowments). However, the United States was the home of infant industry protection at the time, and many of its tariffs were not for revenue reasons. Moreover, O'Rourke (2000) and Lehmann and O'Rourke (2008) show that the positive tariff-growth statistical correlation is not driven primarily by the New World countries.

6. Clemens and Williamson (2001) argue, on the basis of an econometric analysis, that around a third of this growth differential between Asia and Latin America during 1870–1913 can be explained by the differences in tariff autonomy.

7. In 1968, Korea's per capita income was $195 in current dollars, against $4,491 in the United States. In 1969, per capita income was $400 in Brazil, compared with $4,803 in the United States. The income data (see www.nationmasters.com) are from the World Bank and the Central Intelligence Agency.

8. The Nationalist Party's constitution was a copy of the Soviet Communist Party constitution. The second president of Taiwan, China, Chiang Ching-Kuo, who succeeded his father,
Chiang Kai-Shek, was a communist as a young man and studied in the Soviet Communist Party school in Moscow with future leaders of the Chinese Communist Party, including Deng Xiao-ping. Korea also had its share of communist influence. General Park Chung-Hee, who masterminded the Korean economic miracle, was a communist in his younger days and was sentenced to death in 1949 for his involvement in a communist mutiny in the Korean army, although he earned an amnesty by publicly denouncing communism. Many of his lieutenants were also communist in their young days.

9. The absolute amount of FDI going into the developing countries has increased by about 14 times from around $21 billion during 1983–89 to around $297 billion during 2002–07. FDI as a share of gross fixed capital formation in developing countries has risen from around 3.3 percent during the 1980s to 11–12 percent since the second half of the 1990s (partly reflecting the relative decline in investment during this period). The share of developing countries in world FDI has risen from 17.7 percent during 1983–89 to 20.7 percent during 1996–2007, if we exclude China (or from 19.6 to 24.3 percent, if we include China). The data are from various issues of the World Investment Report (UNCTAD various years).

10. Some countries reduced such ceilings substantially. For example, India cut its trade-weighted average tariff from 71 to 32 percent. However, many countries, including India, have fixed them at relatively high levels. For example, Brazil cut its trade-weighted average tariff from 41 to 27 percent, Chile cut it from 35 to 25 percent, and Turkey cut it from 25 to 22 percent (Amsden 2005, 219, table 11.2).

References


Comment on “Industrial Policy: Can We Go Beyond an Unproductive Confrontation?” by Ha-Joon Chang

KAREN R. POLENSKE

The topic for this session is extremely timely, given the global economic and financial crisis, which, I believe, shows that many policy makers got the economic and financial policies desperately wrong. As I was reading Chang’s paper, I wondered whether he was advising President Obama or whether President Obama’s new industrial and policy directives are helping to justify Chang’s propositions concerning industrial policy.

Chang has three objectives in his comprehensive paper. First, he explores both sides of the industrial policy debate, revealing its successes and failures, but still showing his bias toward the proponent side. Second, he critically examines many of the key issues in the debate, including (1) whether targeting particular industries is a desirable industrial policy, (2) whether the state can ever beat the market, (3) implementation issues, (4) performance indicators, emphasizing the importance of exports, and (5) the need to combine free trade, export promotion, and infant industry protection. Third, he discusses the evolution and current changes in global rules of trade and investment. Although I agree with many of his arguments, I hope that analysts will go beyond industrial policy to study carefully the related issues of ecological, energy, environmental, poverty, and regional policies.

First, I review some of the key myths that Chang tries to knock down, referring not only to his paper for this conference but also to his comprehensive book on the topic (Chang 2002) and several of his recent papers. Then I expand the spatial horizon to include China in greater depth than he does, partially based on my 24 years of working in China, recently on the coke-making sector (Polenske 2006), and review the growth rates of East Asian countries. I end by posing some questions relevant to the global economic turmoil and prospects for industrial policies in the future.

Karen R. Polenske is Peter de Florez Professor of Regional Political Economy at the Massachusetts Institute of Technology. She thanks Mr. Ning Wu for preparing the figures for inclusion in this paper.

Annual World Bank Conference on Development Economics 2010, Global
© 2011 The International Bank for Reconstruction and Development / The World Bank
In all the countries he covers, Chang (2002, 2008) notes the important role played not only by targeted industrial policy, but also by education, government support, physical infrastructure, such as roads and railroads, research and development, and skill formation. He extensively discusses Britain, France, Germany, Sweden, and the United States as well as the East Asian region.

He reminds us of the way in which several important early political economists, such as Daniel Raymond, Matthew Carey, and Henry Carey (Matthew’s son) have been “airbrushed out of the history of U.S. economic thought” (Chang 2002, 31). I do not remember those names in the many fascinating economic history books that Professor Alexander Gerschenkron had us read at Harvard University in the 1960s, including his own book *Economic Backwardness in Historical Perspective* (Gerschenkron 1962). I did experience such “airbrushing” in person (on a different topic) when I worked on the eighth edition of *The Theory of Monopolistic Competition* with Professor Edward H. Chamberlin (Chamberlin 1962), who had published his first edition in 1933. Chamberlin proudly produced as many editions as Alfred Marshall’s well-known theory book *Principles of Economics* (Marshall 1920). I am probably the only one who knows that in the eighth edition Chamberlin deleted key passages from the preface to the seventh edition in which he had taken Joan Robinson to task for her book *The Economics of Imperfect Competition* (Robinson 1933). When I questioned this deletion, his simple reply was, “I have won that argument,” and he was not bothered that he was deleting something from a previous preface, which, of course, he could not do in reality, but only in the published eighth edition.

Also, relevant to today’s session, I note that Marshall worked for several years on a second volume, which was to cover foreign trade, money, taxation, trade, and collectivism. It is interesting to speculate on whether such a volume would have influenced the industrial policy debate as much as his texts *Principles of Economics* and *Economics of Industry* (Marshall 1919, 1920) still influence economic analysts. Piore and Sabel (1984) reintroduced us to Marshall’s *Economics of Industry* in their book *The Second Industrial Divide*.

**Chang’s Explosion of Myths**

In his paper, Chang extends and expands on his earlier writings on industrial policies in important ways. Even so, I found it useful to review his 2002 book *Kicking Away the Ladder*, in which he carefully explodes several myths concerning industrial policies and related economic development by providing extensive documentation of growth rates and tariff policies in several key Western countries as well as in Japan and a few other Asian countries. Although I agree with many of his arguments, I raise in the conclusion several unanswered questions.

The main myth examined by Chang is that free trade, rather than labor and capital mobility, is the key to global prosperity. As Chang (2003) says, historically, the current industrialized countries, such as the United States and Britain, “promoted their national industries through tariffs, subsidies, and other measures” rather than
by abandoning interventionist policies as advocated by neoliberals, who emphasized the virtues of small government, laissez-faire policies, and open international trade policies.

Chang (2002, 17) presents a convincing table from Bairoch (1993) in which the manufactured product tariff rates for the United Kingdom and the United States in 1820 are shown to average 45–55 and 35–45, respectively, in percentage of value. For the United States, these rates remained high until 1950, but even then, they were 14 percent—hardly the free trade environment that some analysts seem to feel existed. Up to the 1900s, tariff protection was an important policy tool (more important than today), because of the limited scope of state intervention and because most countries did not have an income tax, did not have a central bank, and owned or regulated few financial institutions and industrial enterprises. In addition to countering those who argue for free trade, Chang (2008) also challenges the proponents of balanced budgets and private ownership.

China

To understand the full relevance of the protection argument, I suggest examining other large countries, such as Argentina (especially the work of Beatriz Nofal), Brazil, India, and China. Given the limitations of time and space and my own 24 years of work in China, I provide a few brief insights into that economy. Figure 1 gives the growth rates of China compared with Japan, the Republic of Korea, and Taiwan, China, from 1978 through 2007. Although all four economies experienced noticeable growth, China stands out as the economy with the highest average annual growth rate, 10 percent during this period (see table 1). Korea and Taiwan, China, follow China with an average growth rate of 6 percent a year. Japan’s average annual growth rate is 2 percent, lower than that of the other economies.

Except China, all of these economies experienced some negative growth during this period. For instance, Korea had an economic downturn in 1980 and again in 1998, with gross domestic product (GDP) growth rates of –1 percent and –7 percent, respectively. Japan had the same situation in 1998 when its GDP growth rate was –2 percent. These negative rates might have been caused by the Asian financial crisis at that time. Taiwan, China, suffered from negative growth in 2001, when GDP declined 2 percent. Compared with these countries, China maintained a strong and steady GDP growth rate, with the lowest rates occurring in 1989 and 1990, when they hit only 4 percent, largely attributable to the tragic incident in Tiananmen Square.

For China, from 1978 to 2007 (the years for which data are relatively readily available), the GDP growth rate averaged 10 percent and was never below 4 percent, according to data from the World Bank. This fact counters many economists who predicted a sluggish rate of growth due to a large number of factors, including lack of private property rights and other institutions and domination of state-owned heavy industries. Jeffrey Sachs and others encouraged China in the 1990s to do away with the trade restrictions, but China, I think fortunately, ignored this and other neoliberal recommendations. China’s GDP growth rates were as high as 14–
FIGURE 1.
Annual GDP Growth Rate in Select Asian Economies, 1978–2007

Source: For China, Japan, and Korea: the World Bank, World Development Indicators online, accessed through Massachusetts Institute of Technology (MIT). For Taiwan, China, International Monetary Fund, World Economic Outlook, 2008 (http://indexmundi.com/taiwan/gdp_real_growth_rate.html).

Note: For China, Japan, and Korea, constant 2000 U.S. dollars. For Taiwan, China, constant new Taiwan dollar (the base year was not specified); data are not available from this source for 1978 and 1979.

TABLE 1. GDP Average Growth Rates for Select Asian Economies, 1978–2007

<table>
<thead>
<tr>
<th>Economy</th>
<th>Average GDP growth rate (%)</th>
<th>1978–2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Korea, Rep. of</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Taiwan, China</td>
<td>6a</td>
<td></td>
</tr>
</tbody>
</table>


15 percent during 1984, 1985, 1992, and 1993, with the lowest rates (4 percent) occurring in 1989 and 1990 (see figure 2).

The Asian financial crisis did have some impact on the Chinese economy, which can be seen in figure 2, with GDP growth staying at 8 percent during 1998 through 2001, a figure that is still significant compared to the performance of other economies. Rosen and Houser (2007, 9) state, “Between 1978 and 2000 the
Chinese economy grew at 9 percent, while energy demand grew at 4 percent. After 2001, economic growth continued apace, but energy demand growth surged to 13 percent a year. It is this fundamental shift in the energy profile of China’s economic growth that has created shortages at home, market volatility abroad, and questions about the sustainability of China’s trajectory.”

As Chang states in a Financial Times op-ed piece (Chang 2008), “If back in the late 1970s it [China] had gone headlong into free trade, rather than a gradual opening-up, it would not have grown at the pace that it has (its economy could have even collapsed in the way that many poorer former Soviet economies did). However, as a result of that strategy, the Chinese economy is today much bigger than what it may have been.”

For the future, I assume that we may not be able to judge well what will work in China and other developing countries, because some critical issues are beginning to emerge, and innovations are taking place not only in the manufacture of products and production of agricultural and mineral products, but also in how governing bodies operate. The environmental and energy crises are affecting what policy makers and the public are willing to do, but the current crisis is not yet reflected in the growth rates. According to Joerss, Woetzel, and Zhang (2009), China has implemented an extensive body of regulations to ensure that it has sufficient and secure energy resources for its continued economic growth, while mitigating the effect of growth on the environment. China is skillfully combining market forces with planned directives.

I do not see everything through these rose-tinted glasses. Members of my research team (and others) have determined that the energy intensity in China has fallen dramatically since 1978—by as much as 67 percent, partially, we find, because of
changes in technology, especially the replacement of obsolete, polluting, energy-inefficient plant and equipment—but that recently the energy intensity in China has turned upward. Rosen and Houser (2007) attribute this mainly to the return to reliance on heavy industry, but from our fieldwork on the coke and steel sectors in China, I think one major important reason may be because the plant managers have not invested in maintenance of their plant and equipment. New equipment that replaces 50-year-old equipment obviously will use less energy and pollute less, but time and again plant managers have indicated to me that they did not purchase the maintenance packages. Money saved in the short term, I think, will come back to haunt China.

Conclusions

I grant the importance of industrial policy in helping to promote economic growth rates in the past and present in many countries. Stopping at that conclusion, however, does not allow us to explore important questions concerning the roles played by the following policies: infrastructure, including its maintenance; education; health; property rights and related institutions; environmental and energy; poverty alleviation; and, of course, given my main area of interest, regional.

References


———. 2003. “Kicking Away the Ladder: The ‘Real’ History of Free Trade.” *Foreign Policy in Focus* (December 30), Washington, DC. http://www.fpif.org/reports/kicking_away_the_ladder_the_real_history_of_free_trade


Social Capital, Institutions, and Development
In recent years a great many scholars have argued that the formation of social capital is the engine of economic progress. Many others have noted, however, that the evidence is mixed. This paper argues that the deep requirement for economic progress is the development of trust among people. Defining social capital in lean terms, namely, as “interpersonal networks,” it shows that when suitably directed, social capital can build and sustain trust; if it is misdirected or if it operates in the wrong sphere, it can hamper economic development and even cause economies to regress. Moreover, if the idea of social capital is to serve a useful purpose in economics, it should be interpreted as interpersonal networks whose members develop and maintain trust in one another to keep their promises by the device of “mutual enforcement” of agreements. Trust is the key to cooperation; “social capital,” when suitably applied, is only a means to creating trust. A natural place to look for the worth of social capital in macroeconomic statistics is in total factor productivity, but doing so implies that total factor productivity is an amalgam of technology and institutions. The paper concludes (in an appendix) by demonstrating how an increase in trust among people would result in an increase in total factor productivity, which is another way of saying that an increase in trust among people would lead to an increase in the economy’s wealth.

The idea of social capital sits awkwardly in contemporary economic thinking. Although it has a powerful, intuitive appeal, social capital has proven hard to track as an economic good. Among other things, it is fiendishly difficult to measure, not because of a recognized paucity of data, but because we do not quite know what
we should be measuring. Comprising different types of relationships and engagements, the components of social capital are many and varied and, in many instances, intangible.

In an early definition, social capital is identified with those “features of social organization, such as trust, norms, and networks that can improve the efficiency of society by facilitating coordinated actions” (Putnam 1993, 167). As a characterization this appears beguiling, but it suffers from a weakness: it encourages us to amalgamate strikingly different objects, namely (and in that order), beliefs, behavioral rules, and forms of capital assets such as interpersonal links (or “networks”), without offering reasons why such an inclusive definition would prove useful for our understanding of the social world.\(^1\) Several authors have subsequently defined social capital even more inclusively, where even attitudes toward others make their appearance: “Social capital generally refers to trust, concern for one’s associates, a willingness to live by the norms of one’s community and to punish those who do not” (Bowles and Gintis 2002, F419).

I argue presently against adopting such a narrow view of the concept.

In developing the economics of what we today allude to as social capital, some authors have focused on a more primitive concept, namely, trust (Dasgupta 1988, 2000, 2007). Others have studied those components of social organization (for example, rotating savings and credit associations, irrigation management systems, credit arrangements, civic associations, and mutual insurance arrangements) that make social capital a productive asset (Levi 1988; Udry 1990, 1994; Besley and Coate 1995; Ostrom 1996; Dasgupta and Serageldin 2000; Grootaert and van Bastelaer 2002). Case studies of the management of local common property resources in poor countries (for example, fisheries, ponds and tanks, forests, grazing lands, and threshing grounds) have offered further insights into the character of those communitarian institutions that enable mutually beneficial courses of action to be undertaken by interested parties (Jodha 1986; Ostrom 1990; Dasgupta and Mäler 1991; Bromley 1992; Baland and Platteau 1996; Ghate, Jodha, and Mukhopadhyay 2008). Yet others have considered a broader sense of the notion by including extended kinship, lobbying organizations, and hierarchical relationships such as those associated with patronage (for example, the Hindu jajmani system, the Sicilian mafia, and street gangs), so that dense networks do not inevitably result in overall economic betterment, at least not in the long run (Gambetta 1993). Moreover, both theory and evidence caution us that communitarian relationships can involve allocations where some of the parties are worse off than they would have been if they had not been locked into the relationships, meaning that, even though no overt coercion is visible, such relationships may be exploitative (Dasgupta 2000, 2008). One can even argue that the theory in question makes precise the sense in which a relationship can be exploitative. In all those accounts, the engagements that rely on social capital occur somewhere between the individual and the state: they are conducted within communities. Indeed, social capital is frequently identified with the workings of civil society (Putnam 1993, 2000).

For some time now, it has seemed to me that in tracking social capital, the most fruitful first step is not to ask what that object might be, but to ask instead a ques-
tion that is faced by any group of people who have agreed on a joint course of action: Under what contexts can they be sanguine that the promises they have made to one another are credible? The question suggests that the fundamental problem facing people who would like to transact with one another is one of trust. The question also points to a lean and usable notion of social capital, one that is shorn of the warm glow that surrounds the notion in the contemporary literature. In what follows, social capital is taken to mean interpersonal networks. I argue, however, that if the idea of social capital is to serve a useful purpose in economics, it should be interpreted as interpersonal networks where members develop and maintain trust in one another to keep their promises by the device of “mutual enforcement” of agreements.

The advantage of such a lean notion of social capital is that it does not prejudge the asset’s quality. Just as a building can remain unused and a wetland can be misused, so can a network remain inactive or be put to use in socially destructive ways. There is nothing good or bad about interpersonal networks: other things being equal, it is the use to which members put a network that determines its quality. But the prior question is one concerning trust.

The Problem of Trust

Trust among people is sometimes taken to be sui generis, arising from motivations that lie somewhere between altruism and self-interest. The Aristotelian conception of trust, which continues to be invoked by political theorists, is hallowed and often shrouded in mystery. In what follows, I use “trust” in the very wide sense of someone entertaining correct expectations regarding someone else’s promises (as in “I trust Mary to do what she promised to do”), including cases where the promises are implicit by virtue of the community’s customary norms (as in “Prakash will not break our trust in him by eating meat when abroad”). So the idea of trust is not limited to the political sphere; it arises in any sphere that involves human interchange.

Imagine that a group of people have discovered a mutually advantageous course of actions. At the grandest level, it could be that citizens see the benefits of adopting a constitution for their country (ranging from the covenant adopted by the Plymouth brethren to the constitution adopted by the Founding Fathers of the United States). At a more local level, the undertaking could be to share the costs and benefits of maintaining a communal resource (irrigation system, grazing field, coastal fishery), construct a jointly usable asset (drainage channel in a watershed), collaborate in political activity (civic engagement, lobbying), do business when the purchase and delivery of goods cannot be synchronized (credit, insurance, wage labor), enter marriage, create a rotating savings and credit association (as in the institution of iddir in Ethiopia), initiate a reciprocal arrangement (“I help you now that you are in need, with the understanding that you will help me when I am in need”), adopt a convention (send one another Christmas cards), create a partnership to produce goods for the market, conduct an instantaneous transaction (purchase something across the counter), and so on. Then there are mutually advantageous courses of action that
involve being civil to one another. They include forms of civic behavior such as not disfiguring public spaces, obeying the law more generally, and respecting the rights of others.

Imagine next that the parties have agreed to share the benefits and costs in a certain way. The agreement could involve some members making side payments to others. Again, at the grandest level the agreement could be a social contract among citizens to observe their constitution. Or it could be a tacit agreement to be civil to one another, such as respecting the rights of others to be heard, to get on with their lives, and so forth. Here we are thinking of agreements over transactions in goods and services. There would be situations where the agreement is based on a take-it-or-leave-it offer that one party makes another (as when a purchaser accepts the terms and conditions in a supermarket). In other contexts, bargaining may be involved (as in a Middle Eastern bazaar). In this paper I do not ask how agreements have been reached, nor look for principles of equity or allocation norms that might have been invoked during negotiation; rather, I ask the question with which we began: Under what circumstances would the parties who have reached agreement trust one another to keep their word?

Because one’s word must be credible if it is to be believed, mere promises are not enough (witness that we caution others, and ourselves too, not to trust people “blindly”). If the parties are to trust one another to keep their promise, matters must be so arranged that (1) at every stage of the agreed course of actions, it is in the interest of each party to plan to keep his or her word if all others plan to keep their word, a condition that ensures that the promises are self-enforcing and (2) at every stage of the agreed course of actions, each party believes that all others will keep their word. If the two conditions are met, a system of beliefs that the agreement will be kept is self-confirming.

Notice that condition 2 on its own is not enough. Beliefs need to be justified. Condition 1 provides the justification. It offers the basis on which everyone could in principle believe that the agreement will be kept. A course of actions, one per party, satisfying condition 1 is a Nash equilibrium (formally, a subgame–perfect equilibrium).

Notice that condition 1 on its own is not enough either. It may be in each agent’s interest to behave opportunistically if everyone believes that everyone else will behave opportunistically. In that case, noncooperation is also a Nash equilibrium, meaning that a set of mutual beliefs that the agreement will not be kept is self-confirming, implying that opportunistic behavior is self-enforcing. Stated formally, a Nash equilibrium is a set of strategies, one per agent, such that no agent has any reason to deviate from his or her course of actions if all other agents pursue their courses of actions. The famous prisoner’s dilemma is a game that has a unique Nash equilibrium in which all parties are worse off than they would have been if they had been able to trust one another to cooperate.

Generally speaking, though, societies harbor more than one Nash equilibrium. Some yield desirable outcomes; others do not. In order to probe the question of which Nash equilibrium can be expected to be reached, if a Nash equilibrium is expected to be reached at all, economists have studied human behaviors that are not
A MATTER OF TRUST: SOCIAL CAPITAL AND ECONOMIC DEVELOPMENT  

Nash equilibriums. The idea is to model the way people form beliefs about the way the world works, the way people behave, and the way they revise their beliefs on the basis of what they observe. The idea is to track the consequences of those patterns of belief formation so as to check whether the model economy moves toward a Nash equilibrium over time or whether it moves about in some fashion or other, but not toward an equilibrium (Evans and Honkapohja 2001).

This research enterprise has yielded a general conclusion. Suppose the economic environment in a certain place harbors more than one Nash equilibrium. Which equilibrium should be expected to be approached, if the economy approaches an equilibrium at all, depends on the beliefs that people held at some point in the past. It also depends on the way people have revised their beliefs on the basis of observations since that past date. This is another way of saying that history matters. Unfortunately, the study of disequilibrium behavior would lengthen this paper greatly. However, a study of equilibrium behavior takes us a long way.

I began by observing that mutual trust is the basis of cooperation and for people to cooperate, conditions 1 and 2 have to be met. I now look for social environments in which conditions 1 and 2 can be met. To do that, it proves useful to classify the social environments in which the promises people make to one another are credible. Four come to mind (Dasgupta 2000, 2007). In the following section I discuss the first two on my list: mutual affection and pro-social disposition. I discuss the latter two social environments in two separate sections because each relies on a distinct form of social infrastructure, namely, external enforcement (an appeal to the rule of law) and mutual enforcement (abiding by social norms). Of the latter two means of attaining cooperation, mutual enforcement is the more fundamental, in that societies need mutual enforcement mechanisms if they are to rely on external enforcement. I argue that mutual enforcement is at the heart of social capital.

**Affection and Pro-Social Dispositions**

I begin with two social environments in which people trust one another to keep their promises.

**Mutual Affection**

Promises are credible if the parties care about one another sufficiently. Innumerable transactions take place only because the people involved care about one another and rationally believe that they care about one another (that is, each knows that the others know that they care about one another, each knows that the others know that each knows that they care about one another, and so on) and thus trust one another to carry out their obligations. Economists model the situation as one where group members have interdependent utilities. The household best exemplifies institutions based on care and affection. As monitoring costs within the household are low (a group of people who cohabit are able to observe and get to know one another), the institution harbors
fewer problems of moral hazard and adverse selection than many other institutions. However, being few in number, members of a household, as a group, are unable to engage in those enterprises that require large numbers of people of varied talents and locations.

**Pro-Social Disposition**

Promises are credible if it is common knowledge that those making the promises are trustworthy or that they reciprocate by keeping their promise if others display trust in them. In the political sphere, that trust could be founded, for example, on a shared conception of citizenship. In less lofty spheres of human interchange, trust could develop from a common understanding of personal integrity. Thus evolutionary psychologists have argued that because of selection pressures that operated among our hunter-gatherer Pleistocene ancestors, we are adapted to have a general disposition to reciprocate. Others have argued that such a disposition is to a greater or lesser extent formed through communal living, role modeling, education, rewards, and punishments and that the process begins at the earliest stages of our lives.\(^3\)

For the purposes here, it is not necessary to choose between the two theories: either would do. In any event, they are not mutually exclusive. Thus, evolutionary psychologists have argued that our capacity to have feelings such as shame, affection, anger, elation, pride, reciprocity, benevolence, and jealousy has emerged under selection pressure (Ehrlich 2000). No doubt culture helps to shape preferences and expectations (thus behavior), which are known to differ widely across societies. But cultural coordinates enable us to identify the locus of points on which shame, affection, anger, etc., are put to work; they do not displace the centrality of those feelings in the human makeup. The thought I am exploring here is that as adults, we not only have a disposition toward behavior such as paying our dues, helping others at some cost to ourselves, and returning a favor, we also practice norms such as punishing people who have hurt us intentionally, higher-order norms such as shunning people who break agreements, on occasion frowning on those who socialize with people who have broken agreements, and so forth. By internalizing specific norms, a person enables the springs of her actions to include them. She therefore feels shame or guilt in violating the norm, and this prevents her from doing so, or at the very least it puts a brake on her, unless she finds overriding considerations. In short, her upbringing ensures that she has a disposition to obey the norm, be it moral or social or personal. When she does violate it, neither guilt nor shame will typically be absent, but frequently she will have rationalized the act. For such a person, making a promise is a commitment, and it is essential for her that others recognize it to be so.

Often enough, the disposition to be honest will be solely toward members of some particular group (clan, neighbors, or ethnic group). This amounts to group loyalty. One may be raised to be suspicious of people from another group and even encouraged to dupe others if and when the occasion arises. Society wastes
resources when people are disposed to be honest only with members of some groups, not others.

The disposition to be trustworthy within both the personal and impersonal spheres exists in varying degrees. When we refrain from breaking the law, it is not always because of a fear of being caught. When an employee in an unorganized sector works overtime, it may simply be a gesture of benevolence, helping out an employer in unexpected need. Recent work in behavioral economics has reaffirmed that benevolence—more generally, pro-social disposition—is not alien to human nature (see, for example, Rabin 1993; Fehr and Fischbacher 2002; Samuelson 2005). The problem is that no society can rely exclusively on such benevolence, for how is one to tell to what extent someone is trustworthy? It is necessary to look elsewhere.

**External Enforcement**

The promises the parties have made to one another to keep to their agreement would be credible if they could devise an institution, or what one may call a *cooperative infrastructure* (Binmore and Dasgupta 1986), in which keeping promises would be in the interest of each party if everyone else were to keep them. The problem therefore is to devise an institution in which keeping to the agreement satisfies conditions 1 and 2. Societies everywhere have constructed solutions to the credibility problem, but in different ways. What all solutions have in common, however, is that those failing to comply with agreements without cause suffer punishment.

It could be that the agreement is translated into an explicit contract and enforced by an established structure of power and authority; that is, an *external enforcer*. However, for an external enforcer to enforce the agreement, it is necessary for breaches to be *verifiable*. Let us imagine that they are verifiable.

By an external enforcer, I imagine, for simplicity of exposition, the state (although it could be the tribal chieftain, a head priest, a warlord, and so forth). The rules governing transactions in the formal marketplace are embodied in the law, meaning that formal markets are supported by a legal structure. The law is enforced by the coercive power of the state. Transactions involve legal contracts backed by an external enforcer, namely, the state.

Why should the parties in question trust the state to carry out its task? After all, the contemporary world has shown that there are states and there are states. The apparatus of the state is controlled by people, so we are faced with an agency problem even there. Simply to invoke an external enforcer for solving the credibility problem will not do; why should the parties trust the state to carry out its tasks in an honest manner? In democracies a possible answer is that the government worries about its reputation. A free and inquisitive press in a democracy helps to sober the government into believing that incompetence or malfeasance will mean an end to its rule come the next election. Knowing that they worry, the parties trust the state to enforce agreements.4
This argument involves a system of interlocking beliefs about one another’s abilities and intentions. Consider that millions of households in many parts of the world trust their government (more or less!) to enforce contracts, because they know that government leaders know that not to enforce contracts efficiently will mean being thrown out of office. In their turn, each side of a contract trusts the other not to renege (again, more or less!) because each knows that the other knows that the government can be trusted to enforce contracts. And so on. Trust among parties to the agreement is maintained by the threat of punishment (a fine, a jail term, dismissal, or whatever) for anyone who breaks a contract. And the parties are confident that the state will honor its agreement to enforce contracts because citizens have coordinated their voting plans. We are in the realm of equilibrium beliefs, held together by their own bootstraps.

Of course, cooperation is not the only possible outcome. Noncooperation can also be held together by its own bootstrap. Each party believes that the others will not keep to the agreement and finds it in its self-interest to break the agreement. At that particular equilibrium, the parties do not trust one another to keep their promises because the external enforcer cannot be trusted to enforce agreements. To ask whether cooperation or noncooperation will prevail is to ask which system of beliefs is adopted by the parties about one another’s intentions. Social systems have multiple equilibriums.

Putnam (1993) offers evidence from Italy to show that if citizens were to invest in social capital (in the sense I am using the term here), they could further their projects and purposes by getting state officials to do their job honestly and efficiently. The underlying mechanism alluded to, however, involves mutual enforcement. I believe that Putnam is right in thinking that social environments involving “mutual enforcement” of agreements are a requirement for those involving “external enforcement.” So I turn to the idea of mutual enforcement.

Mutual Enforcement as a Feature of Social Capital

I begin by imagining that the group does not have access to an external enforcer. Suppose, however, that they expect to face similar transaction opportunities in each period over an indefinite future. Assume also that breach of agreement is observable by all in the group. In such a situation, the parties could be sanguine that their agreement would be kept if it were to be mutually enforced. The basic idea is this: a credible threat by members of the group that stiff sanctions will be imposed on anyone who breaks an agreement will deter everyone from breaking it. The problem is to make the sanctions sufficiently stiff and the threat credible. The solution to the credibility problem in this case is achieved by recourse to social norms of behavior.

Recall that a strategy means a set of conditional actions. Strategies take the following form: “Do P if X happens, but Q otherwise,” “I will do M if she does Y, but N otherwise,” and so forth. A social norm means a strategy that is followed by members of a community. So a social norm is a “rule of behavior.” However, for a
rule of behavior to be a social norm, it must be in the interest of everyone to act in accordance with the rule if all others were to act in accordance with it. Social norms are (Nash) equilibrium rules of behavior.

Notice that a social norm does not insist that everyone must follow the same rule of behavior. In Indian villages, a Brahmin’s role differs considerably from that of someone belonging to a schedule caste. The strategy underlying a social norm can be label specific, be the label caste, class, age, or marital status.

I now show how a reliance on social norms can be the basis of mutual trust among people. For a thorough, technical exposition, see Mailath and Samuelson (2006). Here it pays to study a numerical example.

**Long-Term Relationships**

Imagine that person A has access to some working capital (raw material, say), worth $4,000. To keep things simple, imagine that A is able to borrow the $4,000 from a source that has the means to recover the debt (for example, because there is a credible external enforcer). I want to avoid having to discuss A’s incentives to repay any debt incurred to the external source because I want to study A’s relationship with B, who has the skills to use the working capital worth $4,000 to produce goods worth $8,000 in the market. A does not have those skills. However, A has access to the market, which B does not. A proposes to advance the capital to her, with the understanding that he will sell the goods once B produces them and will share the proceeds with her. If B was not to work for A, she would use her time to produce goods for her home, worth $2,000 to her. In order to get her to accept his offer, A proposes a sharing rule that is hallowed by their tradition: the $8,000 would be used first to compensate both parties fully ($4,000 for A and $2,000 for B), and the remaining $2,000 would then be divided equally between the two. A would receive a total of $5,000 and B, $3,000. Each would gain $1,000 from the arrangement.

B regards the proposal as fair, but is worried about one thing: Why should she trust A not to renege on the agreement by keeping the entire $8,000 for himself?

Imagine that the opportunity for A and B to do business with each other is expected to arise over and over again, say, annually. The time taken for B to produce her output is assumed to be well within a year. Let $t$ denote time ($t = 0,1,2,...$). Although the future benefits from cooperation are important to both A and B, they will typically be less important than present benefits, because, among other things, there is always the chance that one of the parties will not be around in the future to continue the relationship or that circumstances may change in such ways that A does not have access to his capital flow. So we suppose that the two parties discount the future benefits from cooperation at the rate $r$. (In the present example, B’s discount rate does not matter.) It is conventional in economics to interpret $r$ as the opportunity cost of capital. That interpretation does not ring true in the present example. So I want to imagine, as is realistic, that at each date there is a risk that the opportunity to do business will not arise again (for example, because A will lose access to his source of working capital). That risk is $r$.\(^5\)
Provided $r$ is small, the pair could in principle enter a successful long-term relationship, where each year A advances $4,000 to B, sells the goods B has produced for $8,000, and pays her $3,000.

Consider the following rule of behavior A might adopt: (1) begin by advancing $4,000 to B, (2) sell the goods if she produces them during the year, (3) share the proceeds according to the agreement, and (4) continue doing so every year so long as neither party has broken the agreement, but (5) end the relationship permanently the year following the first defection by either party. Similarly, consider the following rule of behavior B might adopt: so long as neither party has reneged on the agreement, work faithfully for A each year, but refuse ever to work for him the year following the first violation of the agreement by either party.

The two rules embody a common idea: begin by cooperating and continue to cooperate so long as neither party has broken his or her word, but withdraw cooperation permanently following the first defection from the agreement by either party. Withdrawal of cooperation is the sanction. This most unforgiving of rules has been christened the “grim strategy,” or simply grim. Next I show that grim is capable of supporting the long-term relationship if $r$ is not too large.

First consider B. Suppose A has adopted grim and B believes that he has. He will advance her the capital at the beginning of year 0. B's best course of actions is clear: keep to the agreement. But suppose she reneges on the agreement. She would lose $1,000 (her share of $3,000 minus the $2,000 she would earn producing home goods), but gain nothing in any future year (remember, A has adopted grim). This means that no matter what B's discount rate is, she could not do better than to adopt grim if A has adopted grim.

The harder piece of reasoning is A's. Suppose B has adopted grim and A believes that she has. If he has advanced the working capital to her, she will have worked faithfully for him in year 0. A now wonders what to do. If he reneges on the agreement, he would make a $4,000 profit ($8,000 minus the $4,000 he could have earned with his capital even if he had not entered into the relationship with B). But since he believes B to have adopted grim, he must also believe that B will retaliate by never working for him again. So, set against a single year's gain of $4,000 is a net loss of $1,000 (the forgone profit from the partnership) every year, starting in year 0. That loss, calculated in year 0, is the sum ($1,000 + 1,000/(1+r) + 1,000/(1+r)^2 + 1,000/(1+r)^3 + ... ad infinitum), which adds up to $1,000(1+r)/r. If $1,000(1+r)/r exceeds $4,000, and it is not in A's interest to break the agreement, which means that he cannot do better than to adopt grim himself. But $1,000(1+r)/r exceeds $4,000 if and only if $r$ is less than one-third a year (or approximately 33.3 percent a year). Therefore, if $r$ is less than one-third, it is in each party's interest to adopt grim if the other party adopts grim. But if both adopt grim, neither would be the first to defect, which implies that the agreement would be kept. Grim can serve as a social norm to maintain a long-term relationship between the patron, A, and the client, B.

Economists have found evidence of grim in social interchanges, but it would appear to be in force mostly where people also have access to formal markets. In the poor world, though, grim is not in evidence. Sanctions are graduated, the first
misdemeanor being met by a small punishment, subsequent ones by a stiffer punishment, persistent ones by a punishment that is stiffer still, and so forth (Ostrom 1996). What is the explanation for this?

Where formal markets and long-term relationships coexist, grim could be expected to be in operation. Grim involves permanent sanctions, which are a device for preventing people from engaging in opportunistic behavior when good, short-term opportunities appear nearby from time to time. But if, as in villages in poor countries, there are few alternatives to long-term relationships, communitarian arrangements would be of high value to all. Adopting grim would be overkill in a world where people discount the future benefits from cooperation at a low rate. For that reason, the norms that are adopted involve less draconian sanctions than grim. A single misdemeanor is interpreted as an error on the part of the defector, or as “testing the water” (to check if others are watching). This is why graduated sanctions are frequently observed.

Here then is the general finding: social norms of behavior are able to sustain cooperation if people care sufficiently about the future benefits of cooperation. The precise terms and conditions will vary across time and place. What is common to them all is that cooperation is mutually enforced; it is not based on external enforcement.

There is, however, a piece of bad news: people could end up not cooperating even if they care a lot about the future benefits of cooperation. To see how, imagine that each party believes that all others will renege on the agreement. It would then be in each one’s interest to renege at once, meaning that there would be no cooperation. Even if $r$ is less than 33.3 percent a year in our numerical example, behavior amounting to noncooperation is also a Nash equilibrium: A does not advance the $4,000 worth of raw material to B because he knows that B will not work for him; she will refuse because of the fear that A will not keep his promise to share the proceeds, a fear that is justified, given that A intends not to share the $8,000 with her once she has produced those goods, and so on. Failure to cooperate could be due simply to an unfortunate pair of self-confirming beliefs, nothing else. No doubt it is mutual suspicion that ruins their chance to cooperate, but the suspicions are internally self-consistent. In short, even when appropriate institutions are in place to enable people to cooperate, they may not do so. Whether they cooperate depends on mutual beliefs, nothing more. I have known this result for many years, but still find it a surprising and disturbing fact about social life.

Could the pair form a partnership if $r$ exceeds 33.3 percent a year? The answer is no. Because grim is totally unforgiving, no other rule would inflict a heavier sanction for a single misdemeanor. The temptation A faces to defect is less if B adopts grim than if she were to adopt any other rule of behavior, which implies that no rule of behavior could support a partnership if $r$ exceeds 33.3 percent a year. Studying grim is useful because it allows us in many examples, such as the present one, to determine the largest value of $r$ for which cooperation is possible.

This tool is useful to explain how a community can skid from cooperation to noncooperation. Ecological stress—caused, for example, by increasing population and prolonged droughts—often results in people fighting over land and natural resources.
Political instability—in the extreme, civil war—could be a reason why both A and B become concerned that A’s source of capital will be destroyed or confiscated. A would now discount the future benefits of cooperation with B at a higher rate. Similarly, if the two are afraid that their government is now more than ever bent on destroying communitarian institutions in order to strengthen its own authority, \( r \) would rise. For whatever reason, if \( r \) were to rise beyond one-third a year, the relationship would break down. Mathematicians call the points at which those switches occur bifurcations. Sociologists call them tipping points. Social norms work only when people have reasons to value the future benefits of cooperation.

Contemporary examples illustrate this. Local institutions have been observed to deteriorate in the unsettled regions of Sub-Saharan Africa. Communal management systems that once protected Sahelian forests from unsustainable use were destroyed by governments keen to establish their authority over rural people. But Sahelian officials had no expertise in forestry, nor did they have the resources to observe who took what from the forests. Many were corrupt. Rural communities were unable to switch from communal governance to governance based on the law: the former was destroyed, and the latter did not really get going. The collective vacuum has had a terrible impact on people whose lives had been built round their forests and woodlands.

Ominously, there are subtler pathways by which societies can tip from a state of mutual trust to one of mutual distrust. The model of the partnership between A and B has shown that when \( r \) is less than one-third a year, both cooperation and noncooperation are equilibrium outcomes. The example tells us that a society could tip over from cooperation to noncooperation owing merely to a change in beliefs. The tipping may have nothing to do with any discernable change in circumstances; the entire shift in behavior could be triggered in people’s minds. The switch could occur quickly and unexpectedly, which is why it would be impossible to predict and why it would cause surprise and dismay. People who woke up in the morning as friends would discover at noon that they are at war with one another. Of course, in practice there are usually cues to be found. False rumors and propaganda create pathways by which people’s beliefs can so alter that they tip a society from one where people trust one another to one where they do not.

The reverse can happen too, but it takes a lot longer. Rebuilding a community that was previously wracked by civil strife involves building trust. Noncooperation does not require as much coordination as cooperation does. Not to cooperate usually means to withdraw. To cooperate, people must not only trust one another to do so, they also must coordinate on a social norm that everyone understands. That is why it is a lot easier to destroy a society than to build it.

**Tying Long-Term Relationships**

Cooperation can be made more robust if the parties tie their agreements. To see how, suppose that in the patron-client relationship we have just studied, the discount rate that A (the patron) uses to value the future benefits of cooperation with B (the client) exceeds one-third a year. We know that for want of trust, the pair are
unable to form a partnership. But now imagine that, in addition to the annual flow of $4,000 worth of working capital, A has access to an annual flow of a different type of working capital, worth $3,000 to him. B does not have the skills to work with that capital, but someone named C does. The time C would need to work A’s capital into a marketable product is worth $1,000 to her. Like B, C does not have access to the market for products. The product can fetch $6,000 in the market, and A is in a position to procure it. A considers approaching C with a proposal to form a partnership: the $6,000 would be used first to compensate the pair; the surplus would then be divided equally between them. Each would enjoy a profit of $1,000 annually. For what values of \( r \) is a partnership between them viable?

Notice that the potential long-term relationship with C is more valuable to A than the relationship he is unable to enter into with B because of a lack of trust. Notice as well that as C’s motivations in the potential relationship are similar to B’s in the previous example, we need not study them again. But we do need to work through A’s reasoning, because the numbers matter. So let us start in year 0. Suppose C has adopted grim. If A advances his capital to her but reneges on the agreement once she has produced the output, he gains $3,000 ($6,000 minus $3,000) that year. Set against it is the $1,000 he would lose every year, starting year 0. That loss, calculated in year 0, is $1,000\((1+r)/r\). If $1,000\((1+r)/r\)$ is less than $3,000, A will renege. If, however, $1,000\((1+r)/r\)$ exceeds $3,000, A can do no better than to adopt grim himself. Since $1,000\((1+r)/r\)$ exceeds $3,000 if and only if \( r \) is less than a half per year (50 percent a year), the pair are able to form a long-term relationship if A’s discount rate is less than 50 percent a year. So suppose \( r \) is less than 50 percent. Then A is able to form a relationship with C, but not with B (\( r \) exceeds one-third).

We are now able to show that, so long as \( r \) is less than 40 percent (or two-fifths) a year, A can form a relationship with B if the three tie the pair of undertakings. Let the proposal be to create both partnerships, but with the understanding that if any party in any year acts opportunistically, both relationships will be terminated. In order to formalize this, let the rule of behavior adopted by B (respectively, C) now read as follows: begin by cooperating with A and C (respectively, B) and continue to cooperate so long as no one has broken the agreement, but cease cooperating with everyone following the first defection by any one in either relationship. Similarly, let the rule of behavior adopted by A now read as follows: begin by cooperating with B and C and continue to cooperate so long as no one has broken the agreement, but cease cooperating with everyone following the first defection by any one in either relationship. Each of the parties has adopted grim once again, but grim here comes with an added sting.

It is easy enough to confirm that B would adopt grim if A and C adopt grim and that C would adopt grim if A and B adopt grim. The interesting exercise is to determine A’s incentives to cooperate if B and C adopt grim. As both clients would terminate their relationship with him if he behaves opportunistically with either, A would defect from both relationships if he defects at all. What remains is to calculate A’s gains and losses if he defects from both relationships in year 0. If he does, he gains $7,000 now ($4,000 from his partnership with B; $3,000 from his partnership with C). Set
against that is the value of all the future benefits from cooperation he will have to forgo. That loss is $2,000(1+r)/r. It follows that A cannot do better than to adopt grim himself if $7,000 is less than $2,000(1+r)/r, which is to say if $r$ is less than 40 percent. As we are supposing that $1/3 < r < 2/5$, we conclude that if relationships are tied, both can be created; whereas if they are kept separate, only the one between A and C can form. The intuition behind the finding is clear. A faces greater temptation to defect from his agreement with B than the one with C, which is why the circumstances under which a relationship could form with B are more restricted than they are for a relationship with C. Tying the two relationships reduces A's temptation to break his relationship with B.

While C does not lose from the move to tie the partnerships, she does not gain either. Only A and B gain. So B has every reason to offer solidarity to C, whom she now regards as a professional comrade. B may even offer a small compensation to C, so as to give her a positive incentive to agree to having the two partnerships tied. In return, C promises to stick by B should A mistreat her. A does not do that, of course, but only because he is smart enough to know that C will break off their relationship if he does.

Further refinements are needed when people who wish to trade with one another are separated by distance. Community responsibility systems in Italy during the twelfth and thirteenth centuries helped people to obtain credit and insurance (Greif 1994). Transgressions by a party were met in a collective way: the group to which the injured party belonged imposed sanctions on the group of which the transgressor was a member. In such arrangements, it is communities, not individuals, who acquire a reputation for honesty. Tying relationships in this manner also creates incentives for members of a peer group to keep an eye on one another. The institution reduces the costs that people incur in keeping an eye on one another.

The drawback of tied relationships among people having different interests is that they require further coordination. If B possesses not only her own skills but those of C as well, and if she has the time to work for A in both ventures, it would be simpler for A to offer both partnerships to B, with the proposal that they be tied. The relationship would involve only A and B, requiring less coordination.

**Culture as Beliefs**

Agreements are kept only because parties expect agreements to be kept. Mutual expectations about “reputation” and “rules of behavior” would seem to require an underlying “thing,” something that would permit the coordination of those optimistic beliefs. But what is that thing? Today we use the term *social capital* to signify that thing. In earlier days it used to be *culture*. But pointing to culture as an explanatory device will not do, because culture itself should be explained.
Basics

We have seen that where incentives are required for cooperation, noncooperation is also a possible outcome. Which state of affairs prevails depends upon mutual beliefs. The theory I am using here does not explain those beliefs; it does identify those that can be rationally held. Rational beliefs are not belied by the unfolding of evidence. As they are self-confirming, rational beliefs offer an anchor for our analysis. Because rational beliefs are not unique, they offer just the kind of flexible anchor we need in order to make sense of societal differences.

In his famous work on the influence of culture on economic development, Weber (1930) takes a community’s culture to be its shared values and dispositions, not just beliefs. Studies as widely cast as Weber’s are difficult to summarize, but the causal mechanism Weber himself would seem to have favored in his work on the Protestant ethic and spirit of capitalism leads from religion, through political culture, to institutions and, so, to economic performance.

Using culture to explain economics has not been popular among social scientists in the postwar period. But there has been a recent revival. The most ambitious appeal to culture to understand differences in economic performance since Weber is Landes (1998), who asks why it is that since the middle of the sixteenth century, countries in Northern Europe managed to race ahead of several others seemingly better placed at the time. No doubt technological progress and its rapid diffusion among populations was the key to that success, but the progress itself needs explaining. The one Landes offers is distinctive, because it gives importance to the evolution (or a lack of it) of different types of attitudes and beliefs in various regions of the world. Landes argues that these differences gave rise to institutional differences (with feedback to attitudes and beliefs), which help to explain why some countries became winners, while others enjoyed a brief period of success before losing to the winners, and yet others merely suffered from atrophy.

Landes offers a historical narrative. An alternative strand of enquiry makes use of statistical evidence, when available. The two strands complement each other. Putnam (1993), Knack and Keefer (1997), and La Porta and others (1997) study cross-sectional data and discover positive links between civic culture (civic engagements, trust) and economic growth, while Granato, Inglehart, and Leblang (1996) study cross-sectional data and find positive links between personal motivation (the desire to advance oneself economically) and economic growth.

The statistical findings should not be given a causal interpretation. The motivation to advance oneself would be expected to depend on one’s expectations (that is, beliefs) regarding the chance that hard work pays off. Parents would be expected to instill personal ambition in their children only if they were sanguine that such ambition would not be thwarted by the social order. And women would not rise beyond their station if they (rationally!) feared retaliation against them for their temerity. Thus even an attitude can be a determined rather than a determining factor. When it
is the former, an observed statistical link between culture and economic progress should be interpreted at most as an equilibrium relationship between two endogenous variables. I use “culture” to denote differences in the beliefs that people hold about one another. Culture in this view is a coordinating device. This line of thinking has been used to explain two contemporary phenomena: the presence of cultural stereotypes (Arrow 1973; Starrett 1976; Coate and Loury 1993) and the extent of tax compliance in a society (Levi 1988; Lindbeck, Nyberg, and Weibull 1999).

**Culture as a Coordinating Mechanism**

Equilibrium beliefs could be the consequence of historical accidents rather than deliberate agreement. Societies that are identical in their innate characteristics (that is, fundamentals) may display very different civic behavior. Similarly, people in one society may harbor cultural stereotypes, even though people in another society possessing the same fundamentals do not harbor them. Culture is not an explanatory variable in either example: it is endogenous in both. Moreover, as my four-way classification of social environments in which people could trust one another to keep their promises suggests, it is not necessary to know someone, even at some steps removed, to form beliefs (even rational beliefs) about his or her intended behavior. Interpersonal networks are certainly necessary if mutually beneficial outcomes are to be identified and the associated agreements reached, but it is not necessary to know each and every fellow citizen to arrive at rational beliefs, at a statistical level, about their intended behavior. Trust is the key to cooperation; what scholars have meant by “social capital” is merely one of the means to creating trust.

Earlier I alluded to disequilibrium beliefs and how social scientists have modeled the way beliefs change over time. We may use those models to explain contemporary cultural differences (differences in rational beliefs) in terms of differences in primitives, such as our material needs, the large-scale ecological landscape, the shared knowledge base, and historical accidents. In such analyses, cultural differences would be correlated with differences in economic performance; they would not be the cause of them.

Different types of variables should be expected to change at different speeds—some slow, some not so slow, yet others fast. Imagine now that certain types of (cultural) beliefs are slow to adapt to changing external circumstances. Since slow variables are to all intents and purposes fixed in the short run, it would not be unreasonable to regard them as parameters for short-run analyses. This is the approximation that social scientists make when they offer cultural explanations for economic performance, for example, the success of Japan in the postwar era (Hayami 1997).

Matters are different in the long run. Individual motivation and beliefs are influenced by values and the practice of norms, and they, in turn, are influenced by the products of society, such as institutions, artifacts, and technologies (Wildavsky 1987). Moreover, any process that ties individual motivations and beliefs to values and norms and thereby to the choices made, and back again, would be expected to be path dependent. There is little evidence, though, that trade and imitation may not lead to convergence in those spheres of culture that have a sizable effect on economic
performance. It is also possible that the effect of a particular component of a people’s culture changes over time even when the culture itself is not changing. The various components of culture are in different degrees complementary to other factors of production. So it is possible for a particular component to lie dormant for decades, even centuries, only to become a potent force when external circumstances are “right.” By the same token, that same component could become ineffective, even dysfunctional, when external circumstances change again. This is why there is no logical flaw in claims such as that Japan’s remarkable economic success in the postwar period has been due in part to some aspects of the nation’s culture, even though those same aspects did not have potency in earlier centuries and may in future even prove to be dysfunctional.

**Networks**

So far I have assumed that interpersonal networks (networks for short) are in place. But networks have to be created. Moreover, searching for others with whom to form networks involves resources (such as time). So it is necessary to study pathways by which networks get formed and the reasons why they get formed.

**Creating Ties**

One may think of networks (social capital) as systems of communication channels for protecting and promoting interpersonal relationships. Interpersonal relationships are a more complex notion than networks, as they are the outcomes of a system of mutual beliefs. But networks cover a wide terrain. They include a unit as tightly woven as a nuclear family and one as extensive as a voluntary organization. We are born into certain networks and enter new ones. So networks are themselves connected to one another. Network connections can also be expressed in terms of channels, although a decision to establish channels that link networks could be a collective one.

An elementary channel connects a pair of individuals directly. However, one can establish indirect links. Person A builds an elementary channel, connecting him to person B, person B builds an elementary channel connecting her to person C, and so forth. A is then connected to C, albeit once removed. Indeed C’s motive for consenting to the establishment of an elementary channel with B could be her desire to be linked to A. And so on.

The clause “personal relationships” is central to the notion of networks. There is also the suggestion that engaging in civic cooperation leads to a heightened disposition to cooperate (Seabright 1993). It amounts to forming personal beliefs about others and one’s own tastes through sampling experiences. But if social engagement fosters trust and cooperation, there would be positive feedback between civic engagement and a disposition to be so engaged. The synergy would be tempered by the fact that the private cost of additional engagements (time) would rise with increasing engagements.
Network Externalities

Installing channels is a way to create trust by getting to know that person, sharing common interests, and so forth. Plausibly, someone’s knowledge of someone else’s character declines with the number of elementary channels separating them, as in perhaps knowing very little personally about a friend of a friend of a friend, knowing rather more about a friend of a friend, and knowing even more about a friend. This creates the necessary tension between the benefits and costs of establishing elementary channels.

But it is possible to be misled by this chain-postulate into thinking that weak ties are not valuable. In fact, they can be very valuable. In a famous study based on interviews with professional and technical workers in a town outside of Cambridge, Massachusetts, Granovetter (1973, 1974) reveals that more than half found their job through a personal connection. The surprise was that the majority of personal connections were mere acquaintances.

Granovetter himself notes that the latter finding should have been expected. The reason that weak ties are especially useful in the search for jobs is that they cover a greater range of links than strong ties. Weak ties connect one to a variety of people and to a wide base of information. However, among rural populations in poor countries, there are more intense ties than weak ties. This narrows the possibilities. But it creates an avenue for migration. One enterprising member of the community moves to the city, perhaps supported by those with whom he has strong ties at home while he searches for work. He is followed by others in a chain-like fashion, as information about job prospects is sent home. Migrant workers may even recommend village relations to their boss, because employing them would reduce moral hazard and adverse-selection problems for the boss. That would explain the still largely anecdotal evidence that city mills often employ disproportionate numbers of workers from the same village. The emotional costs of adaptation to new surroundings would also be lower for later migrants, with the implication that migration in response to new opportunities in the city would be slow to begin with but would pick up strength as costs decline (Carrington, Detragiache, and Vishwanath 1996). Formal evidence of chain migration, though sparse, does exist. Caldwell (1969) confirms its occurrence in Sub-Saharan Africa, and Banerjee (1983) provides evidence from an Indian sample. Chain migration from village to town is observed among children in Karnataka, India, by Iversen (2002) in his study of peer-group emulation as a determining factor in the supply of child labor.

There can also be negative externalities in the creation of channels, such as those within groups that are hostile to one another. One would expect an oversupply of them (they are often neighborhood “arms” races; Gambetta 1993). Be they positive or negative, externalities give rise to collective inefficiency. Positive externalities point to an argument for public subsidy; negative ones point to an argument for investment in institutions such as those whose presence would lower the externalities (taxing” the corresponding activities would be another possibility). Local authorities frequently apply this argument when establishing youth centers, social clubs, and the like.
The Strength of Inherited Networks

Wintrobe (1995) asks why social networks frequently operate along ethnic lines and why they are multipurpose and dense, unlike specialized professional networks—that is, why narrow identities are assumed so frequently along ethnic lines. In answer he observes that exit from and entry into ethnic networks are impossible and suggests that the threat of sanctions by the group prevents children from reneging on their tacit agreement to work within the networks.

Additional forces are also at work. It should not be surprising that the social channels that people bequeath their children in traditional societies frequently amount to ethnic networks (who else is there with whom one can form connections?). Posner (1980) observes in the African context that because monitoring one another’s activities is not costly within village and kin groups, confining networks to them is a means of reducing moral hazard and adverse selection. But although it is true that exit from one’s ethnicity is literally impossible, children may choose not to use the ethnic channels they have inherited. So Wintrobe’s thesis needs to be extended if it is to explain why those particular networks are so active: their mere denseness would probably not suffice. The way to extend the account is to observe first that investment in networks is irreversible. One cannot without cost redirect channels once they have been established, because such investments are inevitably specific to the relationships in question. Moreover, if trust among people begets trust (Seabright 1993, 1997), the cost of maintaining a channel would decline with repeated use (witness that we often take our closest friends and relatives for granted). So using a channel gives rise to an externality over time, much as in “learning by doing” in the field of technology use. The benefits from creating new channels are therefore low if one has inherited a rich network of relationships, which is another way of saying that the cost of not using inherited channels is high. Outside opportunities have to be especially good before one severs inherited links. This explains why individuals maintain so many of the channels inherited from their family and kinship and why norms of conduct pass down the generations. Individuals are, so to speak, locked in from birth.

Sundry Features of Social Capital

Three features of social capital (that is, networks whose members enter into engagements under the discipline of mutual enforcement) deserve special attention. I turn to them here.

Narrow Identities

There is a close link between “social capital” and “social identities.” Activities in networks create bonds, sometimes even affection, among members. Trust develops on the basis of the first two contexts identified under the headings mutual affection and pro-social disposition. Here a person’s social identity is defined by the networks to which she belongs.
How many networks would a person be able to join? It is a truism today that a person’s identity is multidimensional and that people share many of the allegiances associated with them. Social psychologists have noted too that aspects of a person’s identity are fluid and built on the deliberative choices of the person himself and of others (Tajfel and Turner 1986). Advocates of liberal cosmopolitanism tell us to recognize humanity whenever and wherever it occurs, while assuring us that it is deserving of our first allegiance and respect (Nussbaum 1996; Maalouf 2000; Barry 2001; Appiah 2005; Sen 2006). Sen (2006), in particular, argues that individuals have multiple identities, so that claims for special and narrow identities are not warranted, even delusionary. And yet, all over the world we see individuals and groups defining themselves in narrow, exclusive terms and defending them vigorously. Why?

Population heterogeneity is a reason: some like one network, others feel more at home in a different network, and so on. Religious groupings are a prime example. Then there is the “lock-in effect” of inherited networks, which makes it costly for someone to leave the networks into which he was born.

There is a third reason. The advantages of tied relationships suggest that size is an advantage to networks. This means that if a member of a network were to join another network in order to further some of his purposes, the former network would incur a loss by being less robust. Therefore, an increase in any given network’s size inflicts a negative externality on other networks. So networks vie with one another for membership. Dasgupta and Goyal (2009) develop a simple model of individual incentives and network interests to identify circumstances where individuals desire multiple identities but are required by networks to assume narrow identities.11

Networks and Human Capital

In his pioneering work, Coleman (1988) sees social capital as an input in the production of human capital. Establishing networks involves time and effort. Much of the effort is pleasurable, but some is not. Even so, just as academics are paid for what they mostly like doing anyway (as a return on investment in their education), networking would be expected to pay dividends even when maintaining networks is a pleasurable activity.

Burt (1992) finds that, among business firms in the United States, controlling for age, education, and experience, employees who occupy a strategic position in networks are more highly compensated than those who do not. His findings confirm that some of the returns from investment in network creation are captured by the investor. However, because of network externalities, not all the returns can be captured by the investor: when A and B establish a channel linking them, the investment improves both A’s and B’s earnings, but it also improves the earnings of C, who was already linked to B.

Burt’s findings suggest that memberships in networks are a component of human capital. If firms pay employees on the basis of what they contribute to profitability, they would look not only at the conventional human capital that employees bring
with them (for example, health, education, experience, and personality), but also at the personal contacts they possess. It would be informative to untangle networks from the rest of human capital. This could reveal the extent to which returns from network investment are captured by the investor. But measurement problems abound. They may be insurmountable because of the pervasive externalities to which they give rise. Moreover, the way that aggregate production functions are specified affects the way that social capital manifests itself in macroeconomic statistics. In the appendix to this paper, for example, I show that even when there are no network externalities, growth in trust among members of a group of people will display itself in growth in total factor productivity if the aggregate production function is suitably formulated.

**Horizontal versus Vertical Networks**

Putnam (1993, 174) observes a critical difference between horizontal and vertical networks:

A vertical network, no matter how dense and no matter how important to its participants, cannot sustain social trust and cooperation. Vertical flows of information are often less reliable than horizontal flows, in part because the subordinate husbands information as a hedge against exploitation. More importantly, sanctions that support norms of reciprocity against the threat of opportunism are less likely to be imposed upwards and less likely to be acceded to, if imposed. Only a bold or foolhardy subordinate lacking ties of solidarity with peers would seek to punish a superior.

There is a third reason. Imagine a network of people engaged in long-term economic relationships, where relationships are maintained by observing social norms (such as norms of reciprocity). Suppose new economic opportunities arise outside the enclave, say, because markets have developed. Horizontal networks are more likely to consist of members who are similarly placed. If one of the parties discovers better economic opportunities outside the enclave, it is likely that others too will discover better economic opportunities. Both parties would then wish to renegotiate their relationship.

Vertical (or hierarchical) networks are different. Even if the subordinate (for example, the landless laborer) finds a better economic opportunity in the emerging markets, it is possible that the superior (that is, the landlord-creditor) does not; in this case the former would wish to renegotiate, but the latter would not. It is no doubt tempting to invoke the Coase argument that the subordinate would be able to compensate the superior and thus break the traditional arrangement (Coase 1960). But this would require the subordinate to be able to capitalize his future earnings, something typically not possible for people who are subordinates in rural economies in poor countries. The promise to pay in installments is not an appealing avenue open to a subordinate either. He would have to provide collateral. As this could mean leaving his family behind, the worker could understandably find it too costly to move.
Networks and Markets

Networks are personal. Members of networks must have names, personalities, and attributes. Networks are exclusive, not inclusive; otherwise they would not be networks. The terms of trade within a network would be expected to differ from those that prevail across networks. An outsider’s word would not be as good as an insider’s word: names matter.

Networks give rise to “communitarian” institutions. In contrast, markets (at least in their ideal form) involve “anonymous” exchanges (witness the oft-used phrase “My money is as good as yours”). To be sure, the distinction between named and anonymous exchanges is not sharp, and even in a sophisticated market (modern banking), reputation matters (credit rating of the borrower). But the distinction is real. The key point that follows is that the links between markets and communitarian institutions are riddled with externalities. Transactions in one institution have effects that spill over to the other without being taken into account. Externalities introduce a wedge between private and social costs and between private and social benefits. I observe below that some externalities are of a kind that reflects synergy between the two institutions, while others reflect antagonism between them.

All societies rely on a mix of impersonal markets and communitarian institutions. The mix shifts through changing circumstances, as people find ways to circumvent difficulties in realizing mutually beneficial transactions. It pays to study those features of goods and services that influence the mix in question and the hazards that lie in wait while the mix changes as a consequence of the individual and collective choices that are made.

Complementarities

Networks and markets often complement one another. Production and exchange via networks in one commodity can be of vital importance to the functioning of the market in another. As economists have long noted, for example, exchanges within the firm are based on a different type of relationship than those in the marketplace between firms.

But complementarities between networks and markets can be a good deal more subtle. Powell (1990) and Powell and Brantley (1992) find that researchers in rival firms in a competitive environment such as the one that prevails in the biotechnology industry share certain kinds of information among themselves, even while they maintain secrecy over other matters. The balance between disclosure and secrecy is a delicate one, but in any given state of play a common understanding would seem to prevail on the kinds of information members of a network of scientists are expected to disclose, if asked, and the kinds one is expected not even to seek from others. In such an environment, noncooperation would be costly to the individual scientist: if he refuses to share information, or is discovered to have misled others by giving false information about his own findings, he would be denied access to information that others share. There is also evidence that sharing research
findings among scientists in rival firms is not clandestine practice. Managers not only are aware of the practice; they positively encourage their scientists to join the prevailing network. Well-connected scientists are especially valued. The geographic clustering of firms in research-based industries (for example, the Silicon Valley in California, the Golden Triangle in North Carolina, and Silicon Fen around Cambridge, England) is a consequence of the need for such networks. Networks can even be the means by which markets get established (long-distance trade in earlier times). In some cases, they are necessary if markets are to function at all.12

**Crowding Out**

Where networks and markets are substitutes, they are antagonistic. In an oft-quoted passage, Arrow (1974, 33) expresses the view that organizations are a means of achieving the benefits of collective action in situations where the price system fails. This formulation, if interpreted literally, gets the historical chronology backward, but it has an important contemporary resonance: when markets displace communitarian institutions in the production of goods and services, some people suffer unless countermeasures are undertaken by collective means.

Arrow’s observation also has a converse: certain kinds of network can prevent markets from functioning well (Arnott and Stiglitz 1991). Networks can even prevent markets from coming into existence. In such situations networks are a hindrance, not a help, to economic development. They may have served a purpose once, but they are now dysfunctional.

To illustrate, consider the strong kinship ties that are prevalent in traditional societies. Such ties reflect a communal spirit absent from modern urban life and strike an emotional chord among Occidental scholars (Apfell Marglin and Marglin 1990). But there is a functional side to kinship ties: the obligation of members of a kinship to share their good fortune with others in the group offers a way to pool individual risks. The lowlands of Sub-Saharan Africa, for example, are in large measure semiarid, where people face large climatic risks. In contrast, people in the highlands enjoy more reliable rainfall. Lineage groups are powerful in the lowlands. They are less powerful in the highlands, where even private ownership of land is not uncommon (for example, the Kikuyu in Kenya; Bates 1990).

However, there is a negative side to the coin. Kinship obligations dilute personal incentives to invest for prosperity. Even if the social return on investment in an activity is high, the private return can be low: because of kinship obligations, the investor cannot appropriate the returns.13 Insurance markets are superior to communitarian insurance systems because the former, covering a wider terrain of people, are able to pool more risks. However, mutual insurance among members of a community (for example, household, kinship, and village) can be expected to be less fraught with problems of moral hazard and adverse selection than markets (Udry 1990, 1994). This means that, if we view kinship obligations over insurance and credit, respectively, as risk-sharing arrangements and intertemporal consumption-smoothing devices, they are to the good, but they are not all to the good, because their presence lowers the private benefits that
people would enjoy from transacting in insurance and credit markets even when the collective benefits remain high.

It is possible to show that the more dissimilar are those engaged in transactions, the greater are the potential gains from the transactions. This means that, to the extent that communitarian institutions are a dense network of engagements, they are like economic enclaves. But if the institutions act as enclaves, they retard economic development. For example, social impediments to the mobility of labor imply that “talents” are not able to find their ideal locations. This can act as a drag on economic development. The same point can be made about credit, if credit is based on kinship. More generally, resources that should ideally flow across enclaves do not do so. Society then suffers from an inefficient allocation of resources.

Micro-Behavior and Macro-Performance

I turn to the issue of how network activities translate into the macro-performance of economies, discovering that they depend on the way aggregate production functions are specified.

Consider a simple formulation of economywide production possibilities. Let individuals be indexed by \( j \) (\( j = 1, 2, \ldots \)). For simplicity, I consider a single manufactured physical commodity. Let \( K \) denote the economy’s stock of physical capital, and \( L_j \) denote the labor-hours contributed by person \( j \). I do not specify the prevailing system of property rights to the reproducible capital, nor do I describe labor relations, because to do so would be to beg the questions being discussed here. But it is as well to keep in mind that in a well-developed market economy, \( K \) would be dispersed private property, in others \( K \) would be in great measure state owned, in yet others much would be in clumps of communally owned property, and so forth. It is also worth remembering that in market economies labor is wage based, that in subsistence economies “family labor” best approximates the character of labor relations, and that labor cooperatives are not unknown in certain parts of the world.

Let \( h_j \) be the human capital of person \( j \) (years of schooling, health). His effective labor input is then \( h_j L_j \). \( h_j \) is what one may call “traditional human capital” (for the moment I leave aside the networks to which \( j \) belongs). Physical capital is interpreted as “manufactured capital,” comprising items such as factories and buildings, roads and bridges, machines and cables, and so on. In short, I ignore natural capital here.

Human capital is embodied in workers. Given the economy’s knowledge base and institutions (taken here to be the engagements brought about by interpersonal networks), human capital in conjunction with physical capital produces an all-purpose output, \( Y \).

Scale versus Change

Write \( H = \sum_j (h_j L_j) \). \( H \) is aggregate human capital. Now suppose that output possibilities are given by the relationship,

\[
Y = AF(K, H), \ (A > 0),
\]

(1)
in which $F$ is the economy's aggregate production function. $F$ is non-negative and is assumed to be an increasing function of both $K$ and $H$.

In equation 1, $A$ is total factor productivity. It is a combined index of institutional capabilities (including the prevailing system of property rights) and publicly shared knowledge. A macroeconomy characterized by the production function $F$ would produce more if, other things being the same, $A$ were larger (that is, if publicly shared knowledge were greater or institutional capabilities were higher). Of course, the economy would produce more also if, other things being the same, $K$ or $h_j$ or $L_j$ were larger. In short, technological possibilities for transforming the services of physical and human capital into output, when embedded in the prevailing institutional structure of the economy, account for equation 1.

Consider now a scenario where civic cooperation increases in the community: the economy moves from a bad equilibrium system of mutual beliefs to a good one. The increase would make possible a more efficient allocation of resources in production. The question arises, would the increase in cooperation appear as a heightened value of $A$, as an increase in $H$, or as increases in both?\textsuperscript{14}

The answer could seem a priori to depend on the extent to which network externalities are like public goods. It may be thought that if the externalities are confined to small groups (that is, small groups are capable of undertaking cooperative actions on their own, with little effect on others, and do take such actions in the good equilibrium), the improvements in question would be reflected mainly through the $h_j$s of those in the groups engaged in increased cooperation. It may be thought, moreover, that if the externalities are economywide (as in the case of an increase in quasi-voluntary compliance in the economy as a whole owing to an altered set of beliefs, even about members of society one does not personally know), the improvements would be reflected mainly through $A$. In the appendix to this paper, I show that the matter is ambiguous: the effect of an increase in trust on the aggregate production function depends on the way the production function is specified to begin with!

For my purposes the ambiguity does not matter, because either way, the directional changes in macro-performance (though not the magnitude of the changes) would be the same. Other things being equal, an increase in $A$ or in some of the $h_j$s—brought about by whichever of the mechanisms has been considered—would mean an increase in $Y$; an increase in wages, salaries, and profits; and possibly an increase in investment in both physical and human capital. The latter would result in a faster rate of growth in output and consumption and, if a constant proportion of income were spent on health, a more rapid improvement in health as well.\textsuperscript{15}

### Interpreting Cross-Sectional Findings

In his analysis of statistics from the 20 administrative regions of Italy, Putnam (1993) finds civic tradition to be a strong predictor of contemporary economic indicators. He shows that indexes of civic engagement in the early years of this century were highly correlated with employment, income, and infant survival in the early 1970s. He also finds that regional differences in civic engagement can be traced back several centuries
and that controlling for civic traditions, indexes of industrialization and public health have no impact on current civic engagement. As he puts it, the causal link appears to be from civics to economics, not the other way around. How do his findings square with the formulation in equation 1?\textsuperscript{16}

The same sort of question can be asked of even less aggregated data. For example, Narayan and Pritchett (1999) analyze statistics on household expenditure and social engagements in a sample of some 50 villages in Tanzania, discovering that households in villages where there is greater participation in village-level social organizations on average enjoy greater income per head. The authors also provide statistical reasons for concluding that greater communitarian engagements result in higher household expenditure rather than the other way round.

To analyze these findings in terms of my macroeconomic formulation, consider two autarkic communities, labeled by \(i (= 1,2)\). I simplify by assuming that members of a community are identical.\textsuperscript{17} Denote the human capital per person in community \(i\) by \(h_i\). By \(h_i\) I now mean not only the traditional forms of human capital (health and education), but also network capital. I denote by \(L_i\) the number of hours worked by someone in community \(i\), by \(N_i\) the size of \(i\)'s population, and by \(K_i\) the total stock of the physical asset in \(i\). Aggregate output, \(Y_i\), is

\[
y_i = A_i F(K_i, N_i h_i L_i). \quad (2)
\]

I do not specify whether improvements in civic cooperation are reflected in increases in \(A_i\), or in \(h_i\), or in both. I leave that specification open here. It follows that if civic cooperation is greater among people in community 1 than in community 2, we would have \(A_1 > A_2\), or \(h_1 > h_2\), or both. Imagine now that the two communities have the same population size, possess identical amounts of physical capital, and work the same number of hours. Gross national product (GNP) in community 1 would be greater than GNP in community 2 (that is, \(Y_1 > Y_2\)). Someone studying the corresponding empirical data would discover that, controlling for differences in \(K\) and \(N_i h_i L_i\), there is a positive association between a community’s cooperative culture (be it total factor productivity, \(A_i\), or human capital, \(h_i\)) and its mean household income (\(Y_i/N_i\)). This is one way to interpret the findings reported in Narayan and Pritchett (1999).

Consider now a different thought experiment. Imagine that in 1900 the two communities had been identical in all respects but for their cooperative culture, of which community 1 had more (that is, in 1900, \(A_1 > A_2\), or \(h_1 > h_2\), or both). Imagine next that since 1900, both \(A_i\) and \(h_i\) have remained constant. Suppose next that people in both places have followed a simple saving rule: a constant fraction \(s_K (> 0)\) of aggregate output has been invested each year in accumulating physical capital. (For the moment I imagine that net investment in human capital in both communities is nil.)\textsuperscript{18} In order to make the comparison between the communities simple, imagine finally that the communities have remained identical in their demographic features. It is then obvious that in 1970, community 1 would be richer than community 2 in terms of output, wages and salaries, profits, consumption, and wealth.
Notice that I have not had to invoke possible increases in total factor productivity \( (A_i) \) or human capital \( (h_i) \) to explain why a cooperative culture is beneficial. In fact, I have deliberately assumed that neither \( A_i \) nor \( h_i \) changes. It is the scale of total factor productivity and human capital that has done all the work in this analysis of the empirical findings; I have not had to invoke secular improvements in them to explain why a more cooperative society would be expected to perform better economically.\(^{19}\)

The problem with this interpretation of the empirical findings is that it does not say how an increase in trust translates into changes in the variables that make up macro-economic statistics. I merely assume that an increase in trust translates into an increase in human capital, an increase in total factor productivity, or both. Is it possible to say something sharper? Is it possible to identify types of trust that, when they increase, translate into an increase in human capital and the types of trust that, when they increase, lead to an improvement in total factor productivity? It transpires that this is not possible because the role of trust in productivity depends on the specification of the aggregate production function, which is a matter of choice.

In the appendix, I present a simple capital model (that is, shorn of index number problems) in which an increase in trust translates into an increase in total factor productivity, even though an increase in trust in the model involves no externalities whatsoever.

**Network Inefficiencies**

As the communities in the thought experiment just conducted are both autarkic, there is no flow of physical capital from one to the other. This is an economic distortion for the combined communities: the rates of return on investment in physical capital in the two places remain unequal. The source of the distortion is the enclave nature of the two communities, occasioned in the example by an absence of markets linking them. There would be gains to be enjoyed if physical capital could flow from community 2 to community 1.

Autarky is an extreme assumption, but it is not a misleading one. What the model points to is that, to the extent that social capital is exclusive, it inhibits the flow of resources; in this case it impedes a movement of physical capital from one place to the other.\(^{20}\) Put another way, if markets do not function well, capital does not move from community 2 to community 1 to the extent that it ideally should. When social networks within each community block the growth of markets, their presence inhibits economic progress.

**Micro-Behavior Again: Dark Matters**

In this paper social capital has been defined as interpersonal networks where trust is maintained by the mutual enforcement of agreements. There is, however, a dark
side of social capital. Two potential weaknesses of resource allocation mechanisms built on mutual enforcement are easy enough to identify.

**Exclusivity**

Networks are exclusive, not inclusive. This means that “anonymity,” the hallmark of competitive markets, is absent from the operations of networks. When market enthusiasts proclaim that one person’s money is as good as any other person’s in the marketplace, they invoke the anonymity property of markets. In allocation mechanisms governed by networks, however, “names” matter. Transactions are personalized. This, as noted earlier, implies inefficiencies. Resources do not move to their most productive uses.

**Inequalities**

The benefits of cooperation are frequently captured by the more powerful within the network. McKean (1992), for example, discovers that the local elite (usually wealthier households) captures a disproportionate share of the benefits of common property resources, such as coastal fisheries and forest products. However, empirical work has, for the most part, only uncovered inequalities in the distribution of the benefits of cooperative behavior. Such findings are, however, compatible with the possibility that all who cooperate benefit. The reason why social capital continues to radiate a warm glow in the literature is that the examples that have motivated thinking on the subject have been coordination games and the prisoner’s dilemma.

Of the two, the prisoner’s dilemma has received the greatest attention. It remains the favorite workhorse in the literature on social capital. The irony is that the prisoner’s dilemma is an uncommon economic game. Dasgupta and Heal (1979, ch. 3) show that when properly formulated, neither the production of public goods nor the management of local common property resources gives rise to the prisoner’s dilemma (see also Dasgupta 2008). Even the famous Cournot duopoly game does not conform to the prisoner’s dilemma. Below I show that the prisoner’s dilemma has kept scholars from exploring the phenomenon of exploitation in communitarian relationships.

**Exploitation**

I began this paper by considering a group of people who have discovered a mutually beneficial course of actions and have agreed to cooperate by following that course. The premise has been that the agreement benefits all members of the network. I now explore the idea that long-term relationships can be bad for some members. In other words, I explore circumstances where some members of a network are worse off being part of the long-term relationship than they would have been in its absence.

That there can be exploitation in long-term relationships should not be doubted. In Indian villages, access to local common property resources is often restricted to
the privileged (for example, caste Hindus), who are also among the more prosperous landowners. The outcasts (euphemistically called members of schedule castes) are among the poorest of the poor. Rampant inequities exist too in patron-client relationships in agrarian societies.

Inequity per se is not evidence of exploitation. But inequities in, say, patron-client relationships are known to take forms that make it likely that the “client” is worse off in consequence of the relationship than he would have been in its absence. Among contemporary societies, there are many where women remain socially inferior beings, prevented from inheriting assets, obtaining education, and entering choice occupations, all of which exclude them from credit, savings, and insurance markets. But such people would appear to accept the restrictions in their lives as a matter of course, without visible or audible complaint. Why?

Dasgupta and Heal (1979, ch. 3) find that in the production of public goods and the management of local common property resources, a player’s min-max value is smaller than the payoff she receives in a noncooperative equilibrium. (In the prisoner’s dilemma the two coincide, which is what makes the game so very special.) That gap (between the equilibrium payoff and the min-max value) can be exploited so that some members of a network are worse off in a long-term relationship than they would have been if the relationship had not been entered into (Dasgupta 2000, 2008). The basic idea is as follows.

Consider a one-shot game possessing a unique noncooperative equilibrium, but where the min-max value of every member is smaller than his equilibrium payoff. The game is not a prisoner’s dilemma. Let us now imagine that the game is to be repeated indefinitely. Let the agreement among the parties read as saying that one of the members is to receive a payoff per period less than her payoff in equilibrium in the one-shot game, but greater than her min-max value.

Call someone a conformist if she cooperates with those who are conformists but punishes those who are nonconformists. This sounds circular, but it is not, because the social norm I want to study requires all parties in the network to start the process by keeping their agreement. It would then be possible for anyone in any period to determine who is a conformist and who is not. For example, if someone were to break the original agreement, she would be judged to be a nonconformist, in which case the norm would require all parties to punish the nonconformist by forcing her to her min-max value. Moreover, the norm would require that such a dire punishment be inflicted not only on those in violation of the original agreement (first-order violation), but also on those who fail to punish those in violation of the agreement (second-order violation), on those who fail to punish those who fail to punish those in violation of the agreement (third-order violation), and so on indefinitely. This infinite chain makes the threat of punishment for errant behavior credible, because if all others were to conform to the norm, it would not be worth anyone’s while to violate the norm. So long as people do not discount future costs and benefits at too high a rate, keeping one’s agreement would then be mutually enforcing.
Conclusions

Writings on social capital have a warm glow about the concept. The observation that relationships matter for a person’s well-being is no doubt trite, but people writing on social capital have claimed more. They have claimed that social capital is an economically productive asset, a source of much that is good about economic and political relationships.

The original literature claims less though. Some regard social capital as an input in the production of human capital (Coleman 1988), while others regard it as the sort of civic engagement that helps to discipline public officials (Putnam 1993). The subsequent literature goes far beyond those modest claims. Among development economists, social capital is interpreted as communitarian relationships. In countries where the law does not function well, where officials regard the public sphere to be their private domain, and where impersonal markets are often absent, communitarian relationships are what keep people alive, if not well; hence their attraction for many contemporary development economists. But we need to bear counterfactuals in mind. It could be that communitarian relationships prevent impersonal transactions from taking place. Moreover, personal obligations inherited from the past can prevent public officials from acting dispassionately. What appears as corruption in the North could well be social obligation in the South. Similarly, one man’s civic association in the North is another man’s special interest group.

In this paper I have suggested that social capital is best seen as interpersonal networks and that, if the concept is to be useful, attention should be paid to engagements within networks that are subject to mutual enforcement. Economists should assess the worth of social capital by studying what networks are engaged in. Some networks would be found to be progressive, others reactionary, yet others violent. That said, the deep underlying feature of an economy that must be present, if the economy is to progress, is not social capital, but rather the extent to which individuals trust one another. In this paper I have also studied how social capital is a means to creating trust. It is a commonplace to say today that an economy’s performance depends on its institutions. True enough, but institutions do not grow in a vacuum. Their functioning depends on trust. In any given historical situation, which institutions should be run on external enforcement of agreements and which on mutual enforcement is a problem to which there is still no firm answer. Determining the right interplay between interpersonal networks and impersonal public institutions remains the central problem of the social sciences. Mutual trust is the elusive bird that all societies would like to capture.

Appendix

How does an increase or decrease in trust translate into macroeconomic statistics? The numerical example in the text captures a salient point: an increase in trust raises incomes by permitting a more efficient allocation of resources: A’s working capital is put to better use under cooperation, as is B’s labor. Consider now two communities
that are identical in all respects, excepting that in one, people have coordinated at an equilibrium where they trust one another, while in the other, people have coordinated at an equilibrium where they do not trust one another. I show below that the difference between the two economies would be reflected in their total factor productivity, which would be higher in the community where people trust one another than in the one where they do not. Enjoying greater income, individuals in the former economy are able to put aside more of their income to accumulate capital assets, other things being equal, so the economy’s wealth would grow faster. Mutual trust would be interpreted from the statistics as a driver of economic growth.

Consider a timeless, subsistence economy of \(N\) households (\(i,j = 1,2,\ldots,N\)). There is a single perishable capital good, which, in combination with labor, can produce a perishable consumption good. The capital being considered is a form of “working capital” (it does not last beyond one period), and output is consumed entirely. Labor is supplied inelastically. If household \(i\) works with \(K_i\) units of capital, it can produce \(F(K_i)\) units of output. This means that households have the same technology of production at their disposal.

I assume that 
\[F(0) = 0, F'(K_i) > 0,\]
and, because labor is also a factor of production, 
\[F''(K_i) < 0.\]
This means that \(F\) is strictly concave. Let \(Y_i\) be household \(i\)’s output. Aggregate output would then be
\[Y = \sum Y_i.\]  
(A.1)

Imagine that households do not trust one another at all, meaning that they are autarkic. Suppose that \(i\) owns \(K^*_i\) units of capital. Under autarky, household \(i\)’s output is \(F(K^*_i)\), which means that aggregate output of the economy, \(Y\), is
\[Y = \sum Y_i = \sum F(K^*_i).\]  
(A.2)

Let \(K^*_i = \Sigma K^*_i\) be the aggregate capital stock. Suppose that households \(i\) and \(j\) form a long-term relationship, while all others remain autarkic. We may imagine that each year \(i\) and \(j\) have access to \(K^*_i\) and \(K^*_j\) units, respectively, of the capital asset. Suppose that \(K^*_i > K^*_j\). Then the two households would maximize their joint output if each were to work with \((K^*_i + K^*_j)/2\) units of capital. (Hint for proof: \(F\) is the same for all households and is a strictly concave function.) This would involve \(i\) giving \((K^*_i - K^*_j)/2\) units of capital to \(j\), with the understanding that \(j\) would, say, repay by sharing \(j\)’s extra produce in some agreed upon manner (recall the sharing rule investigated in the example of the “putting out” system of cooperation in the main text). Household \(i\)’s (respectively \(j\)’s) output would be
\[F(K^*_i - (K^*_i - K^*_j)/2) = F((K^*_i + K^*_j)/2),\]
respectively, 
\[F(K^*_j + (K^*_i - K^*_j)/2) = F((K^*_i + K^*_j)/2)].\]

By the strict concavity of \(F\), we have
\[2F((K^*_i + K^*_j)/2) > F(K^*_i) + F(K^*_j),\]  
(A.3)

which is why it pays \(i\) and \(j\) to reach an agreement. The incomes \(i\) and \(j\) enjoy are larger because of the agreement, but the incomes enjoyed by all other households...
remain unaffected. It is tempting to say that the human capital of only $i$ and $j$ has increased. But aggregating household production functions into an economywide production function tells an ambiguous story. Below I show that the aggregate production function can be defined so that social capital shows up as total factor productivity, implying an increase in the marginal productivity of every household’s labor input!

If only $i$ and $j$ reach an agreement, the outputs of all other households remain the same, which implies that the agreement between $i$ and $j$ creates no externalities. But because the aggregate output of $i$ and $j$ increases, economywide output increases and so on for all other possible networks that may form. I conclude that all possible networks would gain by sharing their initial endowments of capital equally and splitting the increased output in some agreed upon manner. If the grand coalition of all households were able to form a giant network, each household would work with $K^*/N$ units of capital.

It is simplest to study the effect on aggregate output of network formation by (a) imagining that households are autarkic and (b) varying the distribution of initial endowments. Cooperation within networks of households can then be studied by tracking the effect of the redistribution of initial endowments on aggregate output in a world where households are autarkic.

Notice that aggregate output, $Y$, would be lowest if households were autarkic and the entire capital asset of the economy was owned by one household. By the same token, $Y$ would be at its highest possible level if households were autarkic and each household had inherited $K^*/N$ units of capital. Write $\alpha_i = K^*_i/K^*$. Thus a distribution of initial endowments can be expressed as a vector on the unit simplex of $N$ dimensions—that is,

$$\alpha = (\alpha_1, \ldots, \alpha_i, \ldots, \alpha_N), \quad \text{where } \alpha_i \geq 0 \text{ for all } i \text{ and } \Sigma \alpha_i = 1. \quad (A.4)$$

If $\alpha$ is the vector of endowment shares, household $i$’s endowment is $K^*_i = \alpha_i K^*$. Because $F$ is strictly concave, we know that for all $\alpha$ satisfying equation A.4,

$$\Sigma F(K^*/N) \geq \Sigma F(\alpha_i K^*) \geq F(K^*). \quad (A.5)$$

Define

$$A(\alpha, K^*) = \Sigma F(K^*; \alpha) = \Sigma F(\alpha_i K^*)/F(K^*), \text{ where } \alpha \text{ satisfies equation A.4.} \quad (A.6)$$

Notice that $A$ is a symmetric function. Notice also that $A$’s minimum value is 1 (when $K^*_i = K^*$ for some $i$) and its maximum value is attained when $K^*_i = K^*/N$ for all $i$.

So equation A.6 yields the following:

$$Y = \Sigma F(\alpha_i K^*) = A(\alpha, K^*)F(K^*). \quad (A.7)$$

It follows from equation A.7 that an increase in social capital, keeping $K^*$ fixed, would be reflected in a larger value of total factor productivity, $A$. Because all redistributions of initial endowments that increase equality raise aggregate income when households are autarkic, it is possible to conclude that when networks form, total factor productivity ($A$) increases.
This result came as a surprise. An increase in trust among a group of households in the model has no effect on remaining households. I had expected asset redistributions to be reflected in changes in human capital. The model says otherwise. Thus whether the formation of networks leads to an increase in human capital or to an increase in total factor productivity is ambiguous. Had I begun constructing the aggregate production by modeling trust among households in terms of their human capital, the resulting aggregate production function would have tracked networks to human capital. This means that the specification of aggregate production functions is somewhat arbitrary.

Of course, the ambiguity that arises from that arbitrariness is sharpest if \( F(K_i) = K_i^\beta \), where \( 0 < \beta < 1 \); that is, the production function is Cobb-Douglas. In this case, equation A.6 reduces to the form
\[
A(\alpha) = \sum K_i^{*\beta} / K^{\beta} = \sum \alpha_i^{\beta},
\]
where \( \alpha \) satisfies equation A.4. (A.8)

In the Cobb-Douglas case, it is not possible to tell the difference between an increase in total factor productivity and an increase in the quantity of human capital.

Notes

1. See also Putnam (2000, 19), who writes, “Social capital refers to connections among individuals—social networks and the norms of reciprocity and trustworthiness that arise from them.”
2. I owe this observation to a reviewer of the paper.
3. See Hinde and Groebel (1991), which contains accounts of what is known of the development processes through which people from their infancy acquire pro-social dispositions; for example, by learning to distinguish accidental effects from intentional effects of others' actions. See also Samuelson (2005) for a classification of experiments that were designed to demonstrate the presence of pro-social disposition in a wide variety of cultures.
4. In autocracies the state may fear a rebellion should it misbehave excessively.
5. Formally, \( r \) is the hazard rate at each date that A and B will not meet again.
6. There can be many more equilibriums, characterized by partial compliance. For expositional ease I mostly restrict the discussion to two extreme equilibriums, those that are characterized by noncompliance and full compliance, respectively.
10. Compare this account with Putnam (1993, 168–69): “Mutual trust is lent. Social networks allow trust to become transitive and spread: I trust you, because I trust her and she assures me that she trusts you.”
11. Large networks can experience communication problems, of course. I assume here that those problems become significant only when networks are very large.

12. Even here the role of networks can be expected to diminish as it becomes easier and easier to transmit and access information in the marketplace.

13. Platteau and Hayami (1998) stress this feature of life in the lowlands of Sub-Saharan Africa. They are concerned to account for differences between its economic performance and that of East Asia since the 1960s.

14. As is well known, it would not be possible to separate the two influences if the production function has the Cobb-Douglas form, \( AF(K, H) = AK^aH^b \), where \( a, b > 0 \). In the text I assume that \( F \) is not Cobb-Douglas.

15. In the text I assume implicitly that wage rates, salary rates, and profit rates are monotonically increasing functions of the marginal products of \( L_j, h_j, \) and \( K \), respectively. In a perfectly competitive world, the former three quantities would equal the latter three, respectively.

16. Putnam stresses the importance of civic engagement for making government accountable and responsible.

17. This is a privilege that theorists are able to enjoy to good advantage. By assuming that potentially different entities are identical, we are able to avoid having to “control for differences” in those same entities. The assumption permits us to better understand statistical correlations within multivariate relationships.

18. It can be argued that the extent to which people save for their future is itself an influence of social capital: people would save more if they trusted their institutions to protect their savings. I abstract from such effects because to include them would merely reinforce the argument I am about to offer in the text.

19. For a different perspective from the one I am advocating here, see Solow (1995), who suggested that if social capital is a potent force in economic development, it should find itself reflected in growth in total factor productivity. In the text I have shown that there need be no growth in the \( A_i \)s for social capital to influence economic performance.

20. A similar argument can be advanced as regards labor mobility and credit.

References


“Individual” Social Capital, “Social” Networks, and Their Linkages to Economic Game

MASAHIKO AOKI

This essay makes a conceptual distinction between societal categories, such as social networks, norms, and social status differentiation, on the one hand, and social capital invested by individual agents on the other. The former are institutional phenomena arising as outcomes of social-exchange games in which individual agents invest strategically in social capital. Recent development of neuroscience suggests the future possibility of measuring social payoffs in terms of trade-off relationships with hedonic payoffs. This possibility may suggest that a linked-game approach to social capital and its implications for economic performance are more promising. This approach might make feasible an analysis of socioeconomic networks that is not possible using an economic or sociological approach in isolation. Particularly, this essay suggests a way to apply the linked-game approach to the problem of the tragedy of the commons, which is becoming one of the most acute public issues of our time.

About 10 years ago, Partha Dasgupta and Ismail Serageldin edited a useful volume on social capital based on a years-long project of the World Bank (Dasgupta and Serageldin 1999). Articles included in that volume exhibited varied views on the concept and ways it can be used in economic analysis, and, in spite of the lucidly executed concluding chapter by Dasgupta that impartially summarized the contents and presented his own comprehensive view, the divide on the concept was not reconciled. Particularly troublesome was the fact that the two giants among contemporary economists—Kenneth Arrow and Robert Solow—raised clear objections to the concept of social capital. Arrow (1999) urged “the abandonment of the metaphor of capital and the
term ‘social capital.’” Solow (1999) concurred in being “critical of the concept of social capital and the way it is used.” What were the reasons for their objections? Arrow, while admitting the plausibility of the hypothesis that social networks could affect economic performance, pointed out that “the reward for social interactions is intrinsic” and lacks the important property of capital as saved for future benefit at deliberate sacrifice today. Solow argued that various things normally referred to as social capital, such as “trust, the willingness to cooperate and coordinate, [and] the habit of contributing to a common effort even if no one is watching,” may be referred to as “behavior patterns” and that the concept of capital in the usual sense of which stock and return can be measured may not be appropriate. Another point made was that social networks or behavior patterns may have positive or negative impacts on economic performance and therefore cannot be considered factors of production like physical or human capital.

In this paper, I dare to stand like a dwarf on the shoulders of the giants and try to see the landscape more clearly. That is, I would like to face squarely these problems and explore whether social capital can still be a useful concept in understanding economic performance. I sense first of all that one of the sources of the divide or debate is that proponents of the concept did not make a clear distinction between—or did not clarify the possible relations between—collective phenomena (or institutional phenomena), such as “social networks” or social “behavior patterns,” on the one hand, and individual actions leading to the accumulation of the so-called social capital, on the other. This question may be related to the “micro-macro transformation problem,” in Coleman’s phrase (Coleman 1988, 1990). Or it may correspond to relations between equilibrium of some kind of social game and individual strategic choices that generate the outcome in the game-theoretic sense. Or, more concretely, one may ask, Who owns social capital: society or the individuals composing it? Second, what really makes social capital, if it exists, distinct from physical capital, human capital, public goods, and so on. What is the “social” about? If it is made definitely distinct from these economic entities, how is it related to economic performance?

In the last few decades, economists have been concerned with trust, norms, and the like, the essence of which is lucidly summarized by Dasgupta (1999). In such ventures, economists regard trust and norms endogenously generated through repeated economic transactions (for example, Kandori 1992; Greif 1994). They represent equilibrium patterns of economic behavior by rational agents. Sociologists, like Granovetter (1985), criticize economists for neglecting the fact that social networks embed and thereby constrain or facilitate (rational) economic transactions. Some economists are prompted to measure “social capital” and “trust” and estimate their impacts on economic performance (for example, La Porta and others 1997; Knack and Keefer 1997; Glaeser, Laibson, and Sacerdote 2002). Which way is right? Is social capital an endogenous variable or an exogenous parameter for economic analysis? Is there a third way of looking at relations between economic and social factors?

In this paper, I explore answers to these questions. I first conceptualize a special type of game—a social-exchange game—distinct from the economic-exchange game. In the social-exchange game, agents invest in their own “social capital” as
individual strategies, while social networks or patterns of behavior with certain characteristics emerge as equilibrium among them. Then I link this game with games in the economic or commons domains and see how individual accumulation and use of social capital in the linked games may generate certain economic outcomes not viable if those economic games stand alone. Possible applications of this idea to economic development and social-economic networks in the information-communications industry and the commercialization of life science are mentioned briefly. Then I discuss how the attribution of differentiated social capital among the agents does, or does not, control the rat race in organizational context, quoting an example drawn from the recent financial crisis. Finally, I apply the notion of individually investible social capital to business corporations. I posit that corporate social capital, as distinct from physical and market-specific reputation capital, may be accumulated by individual business corporations through a corporate social responsibility program and examine its potentially significant implications for business strategy and public interests. A final section concludes.

Conceptualizing an Individual’s Social Capital: The Social-Exchange Game Approach

One of the early proponents of the concept of social capital, James Coleman, employed the following interesting analogy to introduce the concept (Coleman 1988, 20):

If A does something for B and trusts B to reciprocate in the future, this establishes an expectation in A and an obligation on the part of B. This obligation can be conceived as a credit slip held by A for performance by B. If A holds a large number of these credit slips, for a number of persons with whom A has relations, then the analogy to financial capital is direct. These credit slips constitute a large body of credit that A can call in if necessary—unless, of course, the placement of trust has been unwise, and these are bad debts that will not be repaid.

Here a game-theoretic notion of social capital as individual assets, independent of their social-systemic consequences, is clearly elucidated. But what kinds of exchanges are involved (What does A do for B?), and what are the payoffs (What is the nature of returns to credit slips?) in those exchanges? How are they different from economic exchanges? Are A and B necessarily placed in symmetric positions to generate social networks characterized by mutual trust and so on?

As mentioned, economists have been endogenizing the notion of trust and social norms as a possible outcome of a reputation game in the economic domain itself. For example, if there is a credible belief among the agents in the community of traders that cheating a trading partner would surely result in the exclusion from further transactions in the community and if they are reasonably far-sighted, then they might refrain from cheating, and honest trading might be supported. But there can be situations where access to some economic goods cannot be denied for technological and other reasons: for example, use of the natural environment. Landscape cannot be hidden, water flow cannot be deflected, and air cannot be denied. But it is intuitive that there may be cases in which social norms could play a role even in such situations
The examination of this kind of situation has important implications for the problem of how to control one of the acute economic problems of today: global warming and other environmental problems. There are other cases of nonexcludability that may not be addressed, at least not entirely, with political commands, legal rules, and the like. To proceed, therefore, I consider first a kind of game that Coleman depicts purely as the social-exchange game and then recouple it with games of economic exchanges to examine implications of the linkage.

Suppose there is a community (group) of agents who mutually interact with each other by means of social symbols (such as words, gestures, gifts, and the like), physical actions (such as help or violence), or the offer of nonmarketable goods (such as valuable information, gossip, and the like) to affect the emotional payoffs of others. Let us call the set of such mutually interactive agents and the sets of their instruments as the domain of social exchanges and their recursive interactions as play of the social-exchange game. A few words may be due to distinguish this class of societal games from economic-exchange games.

First, although exchanges of social symbols (for example, speech acts) may be involved in other types of domains as well, those in the social-exchange domain are distinct because of the nature of the exchanges as well as the objectives of players. Any economic exchange is essentially a contract that cannot be implemented without a mutual agreement reached ex ante between two (or more) specific parties, although it may be unilaterally or bilaterally defaulted. However, social exchanges can be made without an explicit agreement, but with “unspecified obligations of reciprocity” (like gift exchange or mutual help), or as a multilateral understanding within a network of agents directly or indirectly connected with unspecified mutual obligations. The conditions in which the unspecified obligations of reciprocity become credible among the agents concerned are discussed later. For now it may be noted that there is a good body of psychological and experimental studies to indicate that reciprocity is deeply embedded in many social interactions (for example, Rabin 1998; Fehr and Gachter 2000; Benabou and Tirole 2006). Second, the utterance of speech or dispatch of other social symbols in social exchanges may well be generated by the sender’s own direct interests or emotions (for example, appreciation, empathy, envy, jealousy, anger, and so on), but the messages are necessarily intended to have an impact on the object's emotional payoffs, either positively (for example, pride, satisfaction, consolation, retribution, and so on) or negatively (for example, shame, guilt, feeling of exclusion, and so on), the kinds of emotions that Elster (1998) characterizes as “social.” In that sense, they are distinct from mere speech acts or the so-called cheap talk (transmission of information with no direct effect on the other’s payoffs) in the signaling game in the economic domain. I assume that engaging in social exchanges may not necessarily be “cheap” for agents, because it may be costly in time and effort sacrificed, value of the gift, and so on.

Each agent can then derive positive or negative “social” payoffs (emotional payoffs) from the social actions of others that are directed toward him or her, as discussed and experimentally verified by neuroscientists. However, in order to expect continual positive actions from others, the agent must reciprocate positive
actions toward them. If somebody is mean to the agent, he or she may wish to seek revenge from the opponent to stop further malicious actions and avoid “social pain.” Thus the expected social payoff of an agent over time in the social-exchange domain will be conditional on the expected actions of others in response to the agent’s own actions. Thus the agent’s social payoff from the social exchange becomes in reduced form the function of his or her own actions, although implicitly via his or her own belief about the other’s reactions to them. The unit of social payoff of his or her own actions may be measured by the marginal opportunity cost sacrificed in terms of the hedonic payoff associated with material consumption in the economic domain. The present value sum of the agent’s expected social payoffs over time is referred to as his or her social capital. It represents the agent’s expected capacity to derive positive net emotional payoffs over time as well as to derive benefits in other domains. Some scholars use the word “social capital” to refer to intangible collective assets that are held by society as a whole (for example, norms and the educational level of the society), analogous to tangible collective assets (for example, public goods and the commons). As discussed shortly, social capital is distinguished conceptually from social norms and other social categories that evolve as societal outcomes of playing the social-exchange games.

An individual agent’s social capital has two features. First, it is the object of individual investment. It depreciates without effort to sustain it. Thus agents exchange social symbolic actions in such a way that they consider the most fit or desirable in order to increase, as well as to make the best use of, their own social capital, albeit with bounded rationality. I provide concrete examples later, but at this point it is enough to point out that the basic structure of social exchanges indicates their strategic nature, albeit not in an exclusively self-regarding manner. Social exchanges based on reciprocity thus defined are different from altruism. Altruism is a form of unconditional kindness. It derives from the nature of agents’ preferences and, as such, is not directly related to either strategic play or the endogenous rules of societal games.

Second, the individual agent’s social capital depends not only on his or her own actions but also on his or her belief regarding the actions of others, the beliefs of others regarding the agent’s beliefs, and so on. In this sense, the social-exchange game shares the same problem of infinite regress as the psychological games introduced by Geanakoplos, Pearce, and Stacchetti (1989) and applied by Rabin (1993). Recent achievement of epistemic game theory suggests that economic agents need a fair amount of common knowledge to arrive at equilibrium beliefs (Aumann and Brandenburger 1995). If agents are recursively engaged in social exchanges within an informative, homogeneous community, then their actions are more easily known and others’ beliefs are more easily inferred. Namely, in the small community, experiences, information, and inferences are shared, resulting in the sharing of behavioral beliefs. Some standard of social exchanges, that is, norms of reciprocity, may then evolve through practices and customs. Such a standard of behavior, or “a behavioral
strategy that is subscribed to by all” (Dasgupta 1999, 341), may be regarded as rep-
resenting a Nash equilibrium of the psychological game.

The shared behavioral beliefs serve as a guide for the agents to act in a socially
proper way. As a Nash equilibrium, it is not beneficial for agents not to follow
them. It is believed that the failure to comply with the norm will be punished
(“sanctioned” in the traditional sociological terminology) by the loss of social cap-
ital. Such loss may not necessarily be implemented by the extreme of external sanc-
tions by others such as ostracism. If norms are internalized, they are followed even
when others would not observe the violation, because not doing so may create
guilt, shame, and other negative emotional payoffs. Such a moral sense need not be
considered as derived from an abstract supernatural axiom or imposed primarily
by an external authority, such as schools, churches, government, and so on. But it
can be regarded as originating in practices. As Aristotle noted, “Moral goodness
(etike) . . . is the result of habit, from which it has actually got its name, being a
slight modification of the word ethos” (Aristotle 1955, 91). Arrow also notes,
“Internalized feelings of guilt and right are essentially unconscious equivalents of
agreement that represent social decisions” (Arrow 1967, 79).

Linking Economic- and Social-Exchange Games

I have presented a simple argument that the genesis of norms may be regarded as
endogenous outcomes in the social-exchange game. However, they do not neces-
sarily have to be regarded as exogenous constraints on choice in the economic
domain. Economic and social exchanges may be linked, and norms may evolve
through interactions of the two. This is how the concept of individual social cap-
ital may still be relevant to an understanding of economic performance. As an
example to exhibit the basic logic in a simple way, let me introduce a simple para-
ble of the dilemma of the commons. That is, suppose that the commons are eco-
nomically or aesthetically valuable to members of the community, but in order to
keep them valuable in a sustainable manner, collective maintenance efforts are
required. Efforts are costly to the members, so there are potential risks of free
riding. Suppose that it is not technologically feasible to exclude any member
from benefiting from the commons. For example, the remarkable growth of rice
production in the early Edo period (from the seventeenth to the eighteenth cen-
turies) in Japan was largely due to the continual reclamation of land and the
associated development of irrigation systems in rural communities. However, rice
paddies cultivated by the individual member families of the community were
scattered and intermeshed due to the incremental land development by fairly
homogeneous member families, while the irrigation system was such that water
drawn from a canal was successively supplied from one paddy field to the next
using the natural slope (the gravity system).7 In this case, the usual reputation
mechanism to control a member’s moral hazard behavior in the collective main-
tenance efforts may not have been feasible only on the commons-game domain.
Suppose further that it is not feasible either politically or for some other reason to solve the problem by establishing individual property rights on scattered and intermeshed paddies or politically forcing the members to be engaged in the collective efforts against their individual incentives and wills. Suppose, however, that the members of the community are mutually engaged in various social exchanges (for example, mutual help as needed, participation in festivities, socializing during leisure and at work, and so forth), which allow the members to derive emotional (and practical) payoffs, with possible costs in time, effort, resources, psychological burden, and so on. The exclusion of any individual from the social exchanges implies the deprivation of his or her social capital.

Even though exclusion from the use of the commons is not technologically possible, shirking of collective efforts in developing and maintaining the commons may be punishable by excluding the shirker from the benefits of the social-exchange game, for example, ostracism. Indeed, in the case of Japan’s Edo village referred to above, severe ostracism known as mura-hachibu was practiced against serious deviants (literally it implies an 80 percent exclusion from village collective actions except for funeral services and fire fighting). In any case, if the common beliefs prevail to the effect that shirking in the commons domain is punishable in the social-exchange domain, the cooperative participation of the members both in the commons game and the social-exchange game becomes the individual’s best choice. Namely, given the shared beliefs, the cooperative states of play in both games are sustained. Such a standard of cooperative behavior, supported by the common beliefs, is referred to as a social norm.

In this example, the social-exchange game is “linked” to the commons game in such a way that the social capital of each agent created in the former game can be used as an incentive for cooperative behavior in the latter game. In that sense, it may be regarded as a game-theoretic restatement of the notion of “social embeddedness,” which originated in the seminal writing of the economic sociologist Granovetter (1985). My point is that social norms are not imposed from outside economic domains that unilaterally embed and control economic behavior. Rather, the norm, or the “pattern of behavior” in Solow’s terminology, is generated and sustained by agents’ strategies to accumulate and use social capital to achieve their own higher overall payoffs inclusive of the social one. Later, I give some examples in which social networks and social capital endogenously evolve together with certain patterns of economic exchanges.

The linked-game approach may also clarify possibly varied roles of social norms for institutional transition. The endogenous view of social norms, which identifies social norms with an equilibrium outcome of reputation mechanisms in the economic domain itself, implies that norms emerge and disappear with relevant modes of economic transactions. Looking at this from a different angle, one may say that the inherent inertia of social norms is in general detrimental to the emergence of a new mode of economic transaction. Greif’s historical comparative institutional analysis (Greif 1994, 2006) provides one instance in which this is true. The cultural belief among the Maghreb traders that dishonest trading will be
punished by ostracism from their community was not shared by outsiders, so the Maghreb traders failed to expand their activities beyond the community. It is claimed that this was a major reason that they eventually lost competitiveness in long-distance trading in spite of their internal efficiency compared with the Genoese traders who relied on efficiency wage discipline to restrain the morally hazardous behavior of agents recruited from markets.

Did pre-market community norms need to be destroyed prior to market transition and replaced by entirely new market mores? How could the latter emerge? Traditional views, whether those of economists (such as Hicks 1969) or scholars in other social science disciplines, such as economic anthropology (Polanyi 1944; Geertz 1983), draw a sharp dividing line between the market economy and the pre-modern economy. A revisionist view contends, however, that under certain conditions, the rural community bound by cooperative norms could play a positive role in facilitating the gradual transition of pre-modern rural economies to market economies. This becomes possible when community norms can regulate morally hazardous behavior on both insiders and outsiders in the initial transition to mutual market relationships. The complete destruction of rural communities may be neither sufficient nor necessary for the emergence of external market relationships and their eventual integration into the market economy. However, as market relationships evolve, the value of social capital that individual members can accumulate through community relationships declines. Various historical and regional examples of the roles of community norms in the transition to the market economy are given in Aoki and Hayami (2001).

Social networks woven by individual investments in social capital also play a role in the development of new types of industrial organizations in which the creation of information is a vital source of competitiveness. In the evolution of the Silicon Valley clustering of start-up firms in the information technology industry, a fair amount of the sharing of information and cognition took place regarding the general direction of developmental potential (that is, the worldview and their position in it), while the development and design of particular modular elements of potentially innovative systems were firmly encapsulated within individual start-up firms. These two opposing aspects are actually complementary in producing innovation systems under high uncertainty. Tournament-like competition among start-up firms can be more conducive to innovation than research and development in an integrated firm under high uncertainty, because parallel development efforts by multiple firms provide option values (multiple experiments; Baldwin and Clark 2000), as well as extra incentives in the form of tournament-winning premiums (that is, the value of a marginal increase in the probability of winning a tournament due to extra efforts; Aoki and Takizawa 2002). However, in order for such a decentralized approach to be feasible, firms have to share information and vision regarding the future possibilities of the industry—that is, how their specialized modular products can be crucial elements in an evolving, innovative product system. Information sharing in this respect is mediated, or brokered, by venture capitalists, university professors, consultants, angels, and so on in a way reminiscent of academic exchanges (for example, Saxenian 1994; Podolny 2001; Burt 2005). To promote exchanges of useful information, reciprocity and mutual trust are needed to restrain exclusively self-regarding—take,
do not give—kinds of attitudes, while informative brokers and helpful mentors are held in high esteem. Thus the dense social exchanges embedded the Silicon Valley mode of industrial architecture, and the agents who invested in social capital in this network, were able to expect economic and social returns in the form of future business opportunities and social status.

Recently some pessimism has prevailed among the venture capital community in Silicon Valley, however. Even after recovering from the bursting of the dot-com bubble in 2002, venture capital investment did not yield satisfactory rates of return, although an abundant amount of financial resources was invested in venture capital funds prior to the 2008 financial crisis. The rate of initial public offerings has declined dramatically since the crisis. It is not yet clear whether this is primarily the by-product of the financial crisis, the maturity of the information technology industry, which is transforming the industrial structure to a more integrated form, or something else. The older generation of venture capitalists who had operational and engineering knowledge has been retiring and been replaced by a new generation whose expertise is in finance more than in engineering and operations. If this is indeed so, the rate of social capital accumulated by the older generation of venture capitalists may be declining.

The organizational field of the commercialization of life sciences has been evolving somewhat differently, as Powell and others (2005) trace and depict with massive data sets and novel analytical methods. In the 1980s hundreds of small scientific firms were established based on scientific knowledge developed in university labs. Some of them soon matured into dedicated biotech firms. In the early days, however, they did not necessarily own skills in marketing drugs that they invented and often relied on pharmaceutical giants, which lacked a new base of knowledge in the rapidly advancing life sciences. This situation did not lead to the acquisition of dedicated biotech firms by the giants, as it did in the information technology field, but it did lead to complementary collaborations between them not only in marketing but also in development efforts. The collaborative networks evolved to include universities, public research organizations, and small start-up firms, as well as venture capital firms and the National Institutes of Health as funding agencies. Since no single organization had the full range of scientific, developmental, and managerial assets needed to produce new medicines, various types of organizations diversified their connections with others in research collaboration and financial support. Career mobility back and forth between universities and industries became commonplace. The overlapping, multiple bundling of organizations as horizontal networks has been evolving, with a small core of organizations at the center. The organizations in the centric positions are there because they have established reputations as being highly effective in having differentiated connections to diverse partners, not just by virtue of being commercially large. As suggested by the fact that these networks involve public research organizations such as universities and hospitals, investment in social capital, as distinct from market-specific reputation capital, has played a significant role for those organizations to be effective. Also, their relationships with other business organizations are not formally contractual; they are more informal and based on unspecified reciprocal obligations. They tend to reject aggressive business firms.
Social Capital and the Rat Race in the Organizational Context

A norm matters even under a competitive framework of organizational architecture in which workers’ jobs are standardized, mutually isolated, and paid for individually by piece-rate contracts, as Burawoy (1979) tells us in his fascinating field study at a Chicago factory. If the piece-rate system stands alone, it could generate a rat race among workers, triggering rate-cut ratchet effects. However, Burawoy depicts the workers as aiming at individually achieving—“making out” in the workers’ slang—a certain individual target rate. Some are satisfied with 125 percent, while others aim higher. But a ceiling (say, 140 percent) is imposed and well recognized by all members. If someone tries to achieve more than the ceiling, he is socially ostracized (although he might not mind being ostracized), while anyone who cannot attain 100 percent is scorned. The author argues that “making out” cannot be understood simply in terms of achieving greater individual earnings; its rewards include relieving boredom, obtaining social relations and psychological rewards, and restraining excessive competition. The culture of “making out” is generated by workers themselves, but once established, it is experienced as a set of externally imposed shop-floor norms. Those who do not follow, or cannot follow, lose or fail to achieve their social capital.

Under other circumstances, however, misaligned incentive contracts in the organization domain may create destabilizing competition among workers by inducing distorted incentives to invest in social capital to earn prestige. For example, in the process leading up to the 2008 financial crisis, the supposedly best and brightest competed on Wall Street and the City to achieve higher social status as symbolized by competence in designing high-yield derivative products. Financial engineering to hedge risks relies on mathematical tools, such as Ito’s lemma, which can be applied to a calculus involving random Brownian movement of variables (say, securities prices). However, in the practical performance evaluation of financial engineers, the so-called tail risks (extreme events with small probability) are not properly taken into account. If events in the lower tail happen, the loss is borne largely by clients who are less informed, while the financial engineers receive higher rewards when events in the high tail occur. Thus they are induced to gamble. However, a large proportion of the increase in revenue from upper-tail events was not real gains; at least a portion of it should have been set aside as reserves for future risks (that is, considered as costs). Thus risk was endogenously amplified rather than hedged.

The “greed” of financiers is often blamed for its disastrous consequences. What does the greed mean in this case? The competition among financial engineers was not simply an economic desire to pursue the insatiable material payoff, but also a social one in which they strived to exhibit their superior competence and intelligence, resulting in emotional payoffs such as a sense of prestige or a feeling of being superior, awed, envied, and the like. Thus there was no ceiling in this rat race, even though a few extra million dollars might not have mattered much to them in a hedonic sense. The economic game and the social game were linked in such a way that the contractual rules of the economic game provided grounds for the culture of rat race in the social-exchange game, which in turn destabilized the state of play in the
economic game. The primary fault may have lain in the misalignment of incentive contracts, but it was the mind-set of people absorbed in the culture that eventually led to the spectacular failure of major financial service firms.

**Do Share Markets Internalize Corporate Social Capital?**

Is there any point to regarding business corporations as engaged (and as ought to be engaged for unique societal benefits) in exchanges with the society of citizens at-large beyond their own markets and business partners? In posing a question in this way, I set aside matters such as corporate brands embodying accumulated reputations in their own markets (consumers’ expectations regarding the quality of products, after-purchase services, timing of delivery, and the like). Costly signaling (such as advertisement) that would not directly affect the utility of the buyers is also left outside the scope of this discussion for a while (although advertisement may promote the so-called conspicuous consumption that enhances social payoffs). The distinction between market-specific reputation and social capital is sometimes subtle and ambiguous in practice, and it is often complementary, as seen later. I start by asking, Do business corporations accumulate (and ought they accumulate) social capital even outside their own product markets? And, if they do, what are the potential implications of this for their market and nonmarket behavior?

Indeed, corporations and citizens of society at-large may be viewed as directly and informally engaged in social exchanges. If a business corporation pollutes a natural environment or generates health hazards through its economic activities or products, it will be subjected to public reactions beyond its own market in the form of a bad social reputation, public protest, a boycott of its products, and so on, even if those economic activities or products are not illegal within the existing statutory framework. On the other hand, some corporations voluntarily provide resources for social benefits such as environmental protection, poverty reduction, public health, education and scientific research, and so on through their so-called corporate social responsibility (CSR) programs. They provide public goods or voluntarily reduce their external costs. As such, these contributions may not immediately contribute to their profits. In response to social contributions that are costly, however, citizens at-large may ascribe social recognition to corporations that provide these resources, contributing to the accumulation of their corporate social capital. To repeat, this corporate social capital is different from market-specific reputation capital. For example, a tobacco company may enjoy a good reputation among smokers, but its corporate social capital may be thin because its products cause health hazards (although tobacco companies are known to be big spenders on CSR programs to cope with their thin social capital accumulation). Why should business corporations be concerned with their social capital? As Friedman (1970) forcefully argues, is it not that individuals (individual shareholders in the capacity of citizens) should contribute to the provision of social benefits, not business corporations whose objective should be to maximize shareholders’ value?
Those who are committed to the so-called stakeholders-society view of corporate governance may argue that the accumulation of corporate social capital constitutes assets collectively beneficial to the stakeholders of the corporation: for example, the employees who have pride in working for an organization known to have a strong social reputation; environmentally conscious citizens who derive satisfaction from owning “green” stocks in the corporation even if they have to give up some dividend income; the embedding community that can expect sustained social contributions from the corporation; the social entrepreneur who has a “warm glow preference” (Baron 2007) for creating a CSR firm even at a financial cost; and so on. Indeed, in this way corporate social capital may cement corporate assets, both human and financial. I refer to the holders of these orientations as the CSR stakeholders.

Although the stakeholder-oriented view thus appears to have merit, a subtle aspect of corporate social capital may not be completely offensive even to the shareholder-oriented view of corporate governance. If investors try to select their portfolio only from stocks of business corporations engaged in CSR, the orthodox financial theory tells us that their financial performance must be worse because they restrict the universe from which stocks can be picked. However, empirical evidence suggests a possible correlation between expenditures for CSR and stock price performance (for example, Dowell, Hart, and Yeung 2000; King and Lennox 2001; Siegal and Vitaliano 2007; Heal 2005, 2009). Why? Two simple, but plausible, reasons could be that (1) profitable corporations may be more willing to contribute to costly CSR and that (2) CSR may be adopted to attract socially responsible consumers and thus enhance productivity (Baron 2007). The first possibility cannot be dismissed outright. For example, socially responsible investments, which now accounts for well over 10 percent of professionally managed funds in the United States, are not performing worse than other funds, although their portfolios are overweighted by the information technology–related stocks that showed relatively stronger growth from the mid-1990s to the financial crisis of 2008. However, event studies by Dasgupta, Laplante, and Maminge (2001) and Dasgupta and others (2004) find that capital markets react to recognition by the media or government of superior (inferior) environmental performance of corporations in a positively (negatively) correlated manner, suggesting causation from environmental performance to stock prices. Also, Siegal and Vitaliano (2007) find that when profits are treated as endogenous (that is, controlling profits due to the “strategic” CSR), they appear to have no influence on the decision to adopt CSR. So it is impossible to reject the possibility that capital markets internalize the values of corporate social capital to some extent.

The logic of capital pricing involving CSR due to Graff Zivin and Small (2005) and Baron (2007) suggests the following story. Suppose that a contribution of CSR is positively but partially (say, θ percent) reflected in the stock value of a corporation. This implies that for citizens-cum-investors who value the corporate giving more than that proportion, the stock price is virtually discounted. They can contribute to a social cause with less cost (that is, 100 – θ percent less). Therefore, contrary to Friedman’s assertion, they are better off buying the stocks of the CSR firm than making a social contribution as individuals. Therefore, the
presence of CSR corporations can increase aggregate social giving. Although the CSR entrepreneurs (and possibly other CSR stakeholders) bear the remaining cost (that is, 100 – \(\theta\) percent of corporate giving), they can derive benefits not only by making their own contributions but also by expanding the opportunity sets for CSR shareholders by providing an alternative to personal giving. But the story is not yet finished.

Business corporations may also be engaged in the development and commercialization of environmentally friendly technology that not only may be appreciated by citizens as a whole but also potentially contribute to its profits: the CSR characterized as “strategic” by Baron (2007). The development can be costly, but its social value may not necessarily be fully appreciated by the potential buyers of its products alone. For example, the potential buyers of eco-friendly cars may be able to save expenditures for fuel after their purchase, but they may not be willing to bear the full development costs charged in the form of a higher price for the car. Thus the managerial calculus of market-specific reputation capital alone may not immediately convince a business corporation to pursue the costly technological development and its commercialization. However, as public concern with environmental degradation and sustainability of energy supplies rises, failure to do so may damage the accumulation of corporate social capital, while investment in environmentally friendly technology may contribute to its accumulation beyond immediate profits from car sales. The ascription of corporate social capital made possible by technological contributions to the cause of society would help the corporate organization to cohere and enhance its cognitive and financial abilities to continue to develop technology.

The ascription of corporate social capital may in turn amplify the value of market-specific reputation, because it may enhance the beliefs of potential buyers of products regarding their user cost-efficiency, durability, and the like, as well as the product’s symbolic value to them (for example, environmental “conspicuous” consumption). In other words, corporate social capital may serve as a positive signal (analogous to advertisement) and improve the long-term profits net of the development cost. Market-specific reputation capital and corporate social capital can be mutually complementary. When de facto property rights in the global commons are shifting from the corporate sector to the public in general, individual corporations must cope with this substantive institutional change by improving their own technological potential and accumulation of social capital. In this situation, the CSR entrepreneur can be motivated to carry strategic CSR beyond the value-maximizing level (Baron 2007).

Corporate CSR activities, pure and strategic combined, thus can link economic, commons, and social-exchange games between business corporations (and their stakeholders such as CSR entrepreneurs and employees) and concerned citizens. Concerned citizens may engage in those games by attributing corporate social capital to CSR corporations, investing in CSR stocks, as well as being potential buyers of the products of CSR corporations. Business corporations are engaged in these games as social givers as well as potential developers of profit-making, environmentally friendly technology. Corporations can do cognitively more than what the
mere collections of individuals can do. This can be especially true with regard to the development of environmentally friendly and renewable energy technology, which requires innovative entrepreneurial initiative, organization of interdisciplinary inputs of knowledge and efforts, foresight, and patience, which CSR corporations, both small and large, may provide. Thus if the linkages of games as depicted in this paper can indeed evolve, strategies that have not been viable in economic calculations alone may become supportable as societal equilibrium.

**Concluding Remarks**

In this essay, I have made a conceptual distinction between social networks, norms, social status differentiation, and so on, on the one hand, and social capital, on the other. The former are social or institutional phenomena arising as outcomes of social-exchange games in which individual agents invest strategically in social capital. Social capital cannot be immediately measured in a way satisfactory to economists, because it cannot be priced in the domain of an economic game based on contractual agreement. But recent development of neuroscience suggests the future possibility of measuring social payoffs in terms of tradeoff relationships with hedonic payoffs. This possibility may suggest that a linked-game approach to social capital and its implications for economic performance are more promising. This approach might make feasible an analysis of socioeconomic network that is not possible using either an economic or a sociological approach. The former regards norms evolving endogenously only in the economic domain and the latter treats social categories as existing prior to individual economic choices. In effect, both treat social and economic factors in delinked manners. However, I have suggested a useful way to apply the linked-game approach to the problem of the tragedy of the commons, which is becoming one of the most acute public issues of our time.

**Notes**

1. The notion of “unspecified obligations of reciprocity” was first emphasized by Blau (1964) as a distinctive characteristic of social exchanges.
2. For a brief survey, see Lieberman and Eisenberger (2008).
3. The tradability between monetary rewards and emotional payoffs is experimentally confirmed by some recent neuroscience studies founded on the notion of “common neuro-currency” (Montague and Berns 2002). A functional magnetic resonance imaging experiment by Izuma, Saito, and Sadato (2008) finds that the acquisition of one’s good reputation activates rewards-related areas of the brain, notably the striatum, and these areas partially overlap the areas related to monetary rewards. See Fehr and Camerer (2007) for a similar study.
4. Putnam’s social capital comes into being not through individual intentional action, but is said to be “inherited,” with its origins hidden in the mists of the past (Putnam 1993). The existing stock cannot be individually owned. Hayami (2009) articulates a collectivist notion
of social capital as “the structure of informal social relationships conducive to developing cooperation among economic actors with the effects of increasing social product.”

5. An individualist view of social capital is elucidated and measured, and its implications are analyzed, by Glaeser, Laibson, and Sacerdote (2002).

6. Such common knowledge may be provided in the forms of culture as encapsulation of past experiences and some public indicators (cognitive societal categories) that make a certain equilibrium the focal point (Aoki 2010, ch. 4). Also see Arrow (1994) for the importance of social categories in achieving equilibrium.

7. The development of collective irrigation systems had quite different paths in Edo Japan and Yi Korea for historical, political, and geographic reasons. I discuss those differences and their possible implications for evolved social norms in some detail in the appendix to chapter 2 in Aoki (2001).

8. For a good survey of this and discussion of related subjects, see Heal (2005).

References


Social capital is a concept with great intuitive appeal, widely suspected to be a valuable contextual factor that might facilitate cooperative economic behavior and hence desirable economic outcomes. Yet, as well argued in the papers by Professor Dasgupta and Professor Aoki, it is still an elusive concept from an economic perspective.

In order to make progress toward answering the key question of whether social capital is a useful notion in understanding economic performance, further advances in conceptualization, theory, and empirics are necessary. It is very hard to ascertain whether social capital helps economic development because we are not quite clear about what we mean by “social capital.” (As argued in Aoki’s paper, many empirical studies on the effects of social capital on economic outcomes are essentially looking at correlations between endogenous variables.) The paper takes a valuable step in the conceptualization of social capital and hence in the eventual understanding of its possible connections to development.

The paper attempts to provide a psychological and economic micro-foundation to the notion of social capital. This is done in a way that overcomes some important objections raised about the concept by scholars of the stature of Arrow and Solow. In particular, the paper suggests an interesting way to operationalize the notions of both “social” and “capital.”

The central theoretical construct in the argument is the notion of a social-exchange game. A social-exchange game is a game in which some instruments (social symbols, physical actions, and nonmarketable goods) are exchanged to achieve “emotional payoffs.”

Aoki builds on the “credit slips” analogy of sociologist Jim Coleman (1990), in order to argue for a game-theoretic notion of social capital as an individual asset. Social capital, in Aoki’s conception, is an object of individual investment, which is accumulated, owned, and used by individual agents: I do things that make you feel...
good today, with the expectation that you will (somehow) reciprocate in the future. This logic overcomes some of the objections to the concept by ascribing to social capital many of the attributes of financial and other forms of capital.

Having settled the connection to capital, Professor Aoki goes about answering the question of what is “social” about it by distinguishing the exchanges in the social-exchange game domain from others (say, economic) in terms of the nature of the exchanges themselves and the players’ objectives. Regarding the nature of the exchanges, it is argued that social exchanges lead to unspecified obligations of reciprocity, unlike the case in most economic transactions (where presumably the terms of exchange are more specific). With regard to the players’ objectives, the social-exchange game is characterized as one in which actions or statements have a direct impact on the object’s payoffs in terms of emotions such as pride, consolation, satisfaction, shame, guilt, and the like, emotions that have been characterized as “social” by Elster (1998, 2007).

Social capital is, then, conceptualized as an individual asset. It is distinguished from social networks, social norms, and other social categories that evolve as societal outcomes of the play of social-exchange games. These are behavioral patterns, supported by sets of beliefs, that appear “in equilibrium” in game-theoretic terms. These categories are closer to the most conventional use of the term social in economics and in common parlance and are also the types of categories that empirical efforts to study social capital attempt to measure.

Social-exchange domains are different from economic-exchange domains. Economic-exchange domains, in a broad sense, include typical economic situations such as the use of common resources and the trade (or more narrowly economic-exchange) domains. Even though social-exchange and economic-exchange domains are distinct, they are connected by the overlap of some of the players. I want to stress some of, since the degree of potential overlap across domains could be an important factor in understanding how the presence of social networks might affect economic performance.

In a simplified manner, we can think of the effects of social capital on economic performance by comparing economic games in isolation to economic games connected to social-exchange games. This comparison could be the basis for some comparative analysis across cases, in which comparable economic domains are “embedded” in varying social contexts.

If we compare a given economic game played in isolation to the same game played in connection to a social-exchange game, various possibilities arise. One (which is the logic implicit in all the literature extolling the economic benefits of social capital) tells us that strategies that were not available in the isolated economic game become available in the larger game, leading to the possibility of more efficient economic outcomes. The paper illustrates this case with a commons situation, in which excludability or other forms of economic or legal sanctions are not feasible, but in which the social-exchange domain supplies the strategic possibility of social exclusion. Adding such possible punishment strategies can lead to more cooperative behavior in the economic domain. Complementary to that logic, it is also possible that the
presence of the connected social-exchange game could allow the selection of a more cooperative equilibrium, even when that equilibrium is also feasible in the original game. Certainly, it is also possible that the “appearance” of an economic-exchange game might break up a previously cooperative equilibrium in the social-exchange domain (as when families who were getting along well get into long-term fights over economic benefits and responsibilities emerging from an inheritance).

As illustrated by those very brief examples, all these connections can work in various ways. Another example provided in the paper is one in which the scale of the social networks, which might have been sufficient to support efficient economic exchanges for previous technologies (say, high transport costs leading to economic exchanges within small geographic scales), becomes a drag on the possibility of extending exchanges when new technologies calling for larger geographic scales come into place. This illustrates what I believe to be a very important issue in the relation between social networks and aggregate economic outcomes: the match or mismatch between the scale or scope of different domains. I return to this point below.

Another very important point raised by Aoki’s paper is that once we think in terms of the dynamics of belief formation and evolution, the beliefs sustaining behavior in social exchanges (call them social norms) do not evolve independently of the economic exchange domain. (The paper provides some interesting examples of the evolution of social norms in the context of specific incentive systems in economic domains.) This is indeed an intriguing point that brings further into question the notion of social capital as an independent exogenous object that influences economic exchange.

Let us focus on the empirical implications of the theory being developed in the paper. Different polities presumably have different levels of social capital. The question is, where does that cross-case heterogeneity come from? It is very unlikely that the difference lies in different brain structures or other nano-foundations of human behavior; the answer must lie in different socioeconomic environments leading to structures of interaction more likely to induce cooperative behavior (supported by internalized norms) in social exchange. That brings us back to the familiar social science terrain of factors such as proximity (geographic or other), likelihood of repeated face-to-face interaction, and the like. (This, by the way, is an additional two-way connection with economic exchange, since in modern societies who we meet often is linked to economic activities.)

The factors likely to lead to high levels of social capital bring us to another issue highlighted in passing: What is the “reach” or “range” of the social-exchange games referred to here? What is the size of the communities engaged in such games? The answer is possibly a wide range, from a small group such as a family to a larger group such as a professional community or a village. The examples in Aoki’s paper cover various ranges. It seems that the concept applies at various levels, a portability that might reflect on its strength. Nonetheless, one is left wondering to what extent the scales at which social-exchange games apply are relevant for issues of aggregate economic performance and economic development. My understanding at this point is that these social-exchange games are likely to be small-scale games, with partial over-
lap across various networks. Whether the type of cooperation relevant for economic development can be supported by the interactions of these various multilayered networks is a question that requires deeper analysis than the one I can provide in these rushed remarks. The remaining comments provide some hints about connecting local social interactions to broad societal outcomes, with focus on a specific domain of articulation.

**Incorporating the Polity Domain**

It would be useful in further developments of the theory suggested here to incorporate another domain that might be important for the questions of economic performance at the macro level in modern societies. That is the *polity domain*, a domain in which the exchanges are political and in which the outcome is public policy.4

It is by now widely recognized that development outcomes are contingent on the content and quality of various public policies. For instance, figure 1 (reproduced from Machado, Scartascini, and Tommasi 2009) shows that an index of the quality of public policies relates to faster growth of total factor productivity (TFP) in the last several decades in a cross section of countries, after controlling for relevant factors such as the level of gross domestic product (GDP).

**FIGURE 1.**

The Quality of Public Policies and TFP Growth

![Graph showing the relationship between the quality of public policies and TFP growth.](source: Machado, Scartascini, and Tommasi 2009.)
So policies are an important factor for economic outcomes. What does social capital have to do with that? The general logic is illustrated in figure 2. The second arrow in the figure is the material of a vast political economy literature establishing the connection between various features of political games and the impact through policies and their properties (such as credibility) on various economic outcomes of interest. The relevant question for the purposes of today’s discussion is the first arrow; that is, the extent to which the type of social exchanges analyzed in Aoki’s paper affect the political game and, through that, the resulting policies and economic outcomes. It is worth noting in passing that the arrow goes both ways, since often some forms of social exchange could emanate from the political-exchange domain (as when long-time political representatives establish personal relationships) or from the policies generated in that domain (when recipients of a particular public program meet repeatedly and establish personal relationships).

Even if we focus just on the direction from the social-exchange game to the political-exchange game, there are various possible connections. Indeed, there is a literature in political science, similar to the one in economics referred to in the papers by Dasgupta and Aoki, which relates social capital to various democratic outcomes. The issue is an ample one, with at least as many ramifications and issues as the study of social capital and economic outcomes. For the sake of brevity, I illustrate the possible connections with two examples of how “local” (small-N) social-exchange game interactions can affect the grand (aggregate) political and policy outcomes. One example focuses on the top of the political game (the relationships among members of the elite), and the other focuses on the bottom of it (clientelistic networks within which some public programs are distributed).

In various countries, including some in Latin America, economic, political, and intellectual elites are parts of dense social networks, abundant in social capital, in many ways resembling the social networks that Aoki’s paper describes for information technology industries. A very integrated elite might have negative effects, if that potential source of intra-elite cooperation is the source of exclusionary politics and policies, as emphasized in the work of Acemoglu and Robinson (2005). Yet, under some conditions, such integrated elites with abundant social capital permit better political exchanges and provide the foundations for better long-term policies.

Some of the most successful economic policies we have seen in Latin America have been in societies where the close ties among elites (and their informal networks) allow for more effective cooperation and long-term horizons. Chile is by any standard the most successful Latin American country in the last couple of decades,
judged by the consistency of its economic policies (leading to the highest growth in the region and poverty reduction). Not only have policy-induced outcomes in Chile been outstanding, but the Chilean political system and way of making policy have been important reasons behind the quality and effectiveness of its policies (Aninat and others 2008). Various features of the Chilean political system have been identified as leading to a more coherent and flexible policy-making regime. One of the factors behind this consensual style of policy making has been a characteristic of Chilean society that, even though it could be problematic for some issues, has facilitated the intertemporal cooperation that has led to successful development outcomes. This characteristic is the fact that its elites are fairly integrated. Political leaders, business leaders, and intellectual leaders are by and large members of the same circles, are educated in the same schools, and have long-term bonds. That leads to a consensual style of interaction between economic and political actors that is rare in the region. The density of business-state networks in Chile is conducive to intertemporal cooperation and to the establishment of patterns of long-term systematic consultation, giving coherence and consistency as well as a long-term orientation to Chilean economic policies (Silva 1997).

Another case that was for many years the regional example (and exception) of successful economic policies was Colombia, a country also characterized (at least until recently) by a closely knit political, business, and technocratic elite. Thorp and Durand (1997) characterize the Colombian elite as a “system in which its members all speak to and know each other. The closeness of the elite, even across the party divide, also allows experts to join rival-party administrations with remarkable freedom, which supports quality and continuity. This results in tight networks and security of access. The private sector feels that it has excellent access to the public sector” (Thorp and Durand 1997, 221). Thorp and Durand (1997, 224) also refer to the “horizontal elite-level characteristics of cooperation and reciprocity.”

A counterexample to the success of Chile (and Colombia) is its next-door neighbor, Argentina, a country with several cultural similarities but with a meek policy performance. Argentine policies have been quite volatile, changing frequently with minor political winds, poorly coordinated across policy areas, poorly implemented, and generally inefficient. The foundations of such policy performance lie in the non-cooperative policy-making environment in which they have been produced. One characteristic of Argentine society that gives rise to such a poor policy-making environment is the inability of elites to generate a consensus on a direction for the country. One of the reasons for this inability is the heterogeneity and incongruence of the Argentine elite (Spiller and Tommasi 2007).

Coming now to a second example of a mechanism connecting (local) social capital to the overall workings of the polity, I briefly discuss the relationship between social capital and clientelism. Clientelism is a form of political exchange in which some personalized material benefits (such as excludable and scarce social programs) are exchanged for votes or other forms of political support (such as participation in rallies). Clientelistic “local” practices could have a negative impact on policies and economic outcomes. Clientelism leads to lower incentives to provide public goods at the local level (Wantchekon 2003 and references therein) and, by
weakening programmatic linkages between citizens and politicians, lowers the quality of national public policies, leading to worse economic outcomes (IDB 2005; Stein and Tommasi 2007).

Even though clientelism is defined as a *vertical* exchange between a patron and a client, it is often embedded in dense social networks of the type emphasized in theories of social capital, which are characterized by long-term, person-to-person interactions in relatively localized geographic spaces (Stokes 2007). Some cross-sectional studies within a given country have found that clientelistic political exchanges are more common in small backward towns, which are characterized by high levels of social capital of the type emphasized in this paper, while other forms of politician-citizen linkages are prevalent in the more anonymous larger-N modern cities, where more impersonal modes of interaction are prevalent (Stokes 2006). As a matter of fact, dense networks of social exchange are likely to provide the conditions for solving one of the key difficulties in clientelistic exchanges: monitoring of the vote. What prevents a “client” from receiving the goods and then voting for whomever he or she pleases? Clientelistic parties depend on their insertion in social networks, especially in low-income voters’ neighborhoods (Auyero 2000). In the words of Stokes (2007, 612–13), clientelistic relationships are ongoing, and the dyad is embedded in a social network: “Networks provide information about their members to other members: we know whether our neighbor or co-worker votes or abstains, voices support for one party or another.” Clientelistic parties use operatives who are embedded in these networks and are “walking encyclopedias of local knowledge.” This local knowledge allows them to make informed guesses about whether a voter to whom the party gave goods or employment actually followed through and supported the party or defected to another. Networks allow clientelist parties to sidestep the secret ballot. The party can then use this information to reward the voter who has cooperated and punish the voter who has defected—it can hold the voter accountable for his or her vote. In contrast to the kind of accountability celebrated in democratic theory, this is “perverse accountability,” in which *parties* hold *voters* accountable for their actions (Stokes 2005).

Even the vertical exchange between political patrons and clients seems to share the psychological micro-foundations stressed in Aoki’s paper. Ethnographic studies of clientelistic exchanges in Latin America and other developing regions emphasize the presence of strong notions of reciprocity in such exchanges. As a matter of fact, one of the first waves of academic work on clientelism, inspired in anthropology and sociology, emphasized moral sentiments different from self-interest as the basic explanation of clientelism. The norm of reciprocity seems to undergird motives for returning benefits beyond what can be accounted for just in terms of self-interest and observability of moves.

**Notes**

1. Aoki (2001) is one of the founding works, together with Calvert (1995a, 1995b), and Greif (1994, 2006), on the very influential view of institutions as equilibriums (a point also
made by Dasgupta in this volume). In particular, Aoki (2001) emphasizes the bounded rationality foundations of institutions, where the common knowledge demanded by game theoretic equilibriums is provided in the form of norms and culture as encapsulation of past experiences. He specifically refers to the “summary” or “information compression” nature of institutions, which he characterizes as self-sustaining systems of shared beliefs about a salient way in which the game is repeatedly played.

2. See Aoki (2001, 21) for the distinction between domains and games.

3. Aoki refers to agents “recursively engaged in social exchanges” within small and homogeneous communities as the cases in which their actions could be more easily known and others’ beliefs more easily inferred, conditions that facilitate the positive economic effects of social capital.

4. The polity domain is studied in chapter 6 of Aoki (2001), where the main features of state institutions are interpreted, in line with the general approach of that book, as stable equilibriums in the polity domain. Those valuable insights notwithstanding, in these comments I suggest a more disaggregated (a more micro) look at the polity domain.


6. These closely knit elite structures also have their drawbacks. The system in Colombia (until recently) has also been characterized as exclusionary, as reflected in the fact that social policies have not been as solid as economic policies. Also, the connection of the elites to the majority of the population have often been characterized as vertical clientelistic networks, a point to which I refer in the remainder of these comments.

7. This is consistent with an assertion of Dasgupta in his paper in this volume, referring to economic exchanges: “It could be that communitarian relationships prevent impersonal transactions from taking place.”


9. Lemarchand (1977) and Scott (1972), cited in Stokes (2007) and in Gallego and Raci-borski (2008). For instance, Lemarchand emphasizes the relationship between clientelism and ethnicity, stating that clientelism extends the perceptions of mutual interest and cultural affinities beyond the realm of primordial loyalties. These exchanges seem to share the characteristic of general and unspecified obligations highlighted by Aoki; Scott (1972) refers to the client reciprocating “by offering general support and assistance.”

References


Financial Crisis and Regulation
Reform of the Global Regulatory System: Perspectives of East Asia’s Emerging Economies

YUNG CHUL PARK

The 2008 global economic crisis has corroborated that strong economic fundamentals and a rational macroeconomic policy framework may not be enough to spare emerging economies from liquidity and currency crises unless they are able to mitigate the consequences of maturity and currency mismatches in the balance sheets of banks and other financial institutions. The mismatching of maturity was one of the major causes of the financial collapse in advanced countries in the run-up to the 2008 crisis. Unlike in advanced economies, maturity mismatch in emerging economies is often accompanied by currency mismatch and a currency crisis when global liquidity evaporates. This is because emerging economies cannot borrow in their own currencies from international financial markets. This disadvantage constitutes a premium on emerging economies’ external borrowing, which means that there is no level playing field for financial institutions from emerging economies engaged in international financial intermediation. Unless institutional arrangements such as a global liquidity safety net, which is on the G-20 reform agenda for global liquidity assistance, are established, emerging economies may not be able to prevent or better manage liquidity and currency crises in the future and hence may have to reexamine the costs and benefits of integrating into the global financial system.

The current global economic crisis has exposed many structural weaknesses in the financial systems of advanced countries. Since the United States and other developed countries dominate international financial intermediation, these weaknesses have been ingrained in and reflected the frailties of the global financial system. The crisis has zeroed in on the efficiency of self-regulation of the market, which has been the cornerstone of Western regulatory philosophy. The widespread misgivings about the

Yung Chul Park is distinguished professor in the Division of International Studies, Korea University.

Annual World Bank Conference on Development Economics 2010, Global
© 2011 The International Bank for Reconstruction and Development / The World Bank
role of the market have given way to tighter regulation of financial institutions and markets. It should come as no surprise that just about every aspect of finance has come under scrutiny for reform—to name a few, issuance and trading of various structured derivative products, shadow banking, and the activities of multinational financial conglomerates headquartered in different jurisdictions.

Some of the conspicuous structural deficiencies were the failings of the international regulatory system. According to the Financial Services Authority (FSA 2009), the global regulatory system failed to identify growing macro- and microeconomic-prudential risks that called for policy responses on the part of national policy and regulatory authorities. It failed to enforce regulatory standards set by standard-setting bodies that were agreed internationally. When the current crisis broke, it failed to organize global efforts to manage the crisis in a coordinated manner. The deepening of the crisis has underscored the urgency of global regulatory reform. Many international forums and international financial institutions have responded to the urgency by proposing an extensive array of global financial reforms. Among these forums, the G-20 has been at the forefront in leading reforms at the national and global levels.

Unlike in previous financial crises, advanced countries have suffered as much as emerging economies from this crisis, largely because it started in the United States and then spread to all parts of the global economy. For this reason, advanced countries have taken the initiative in reforming the global as well as their own regulatory systems. Because of this focus, most of the reports that have been issued do not specifically address the regulatory issues confronting emerging and developing economies. For their part, emerging economies, preoccupied with domestic economic issues, have not shown much interest in global regulatory reform. This is unfortunate because they have not been immune to the collateral damage of the crisis, and the proposed reforms will have far-reaching implications for the restructuring and management of their regulatory systems as well.

During the Asian crisis of 1997–98, East Asia’s emerging economies except for China sustained heavy losses in output and employment, as they were unable to fend off what was essentially a capital account crisis. In the aftermath of the Asian crisis, they were subject to a large array of financial and corporate sector reforms, all of which were designed to prevent and improve the management of future crises. Ten years later, before the eruption of the current crisis, there was general consensus that East Asia’s emerging economies had made great strides in their reform (see, for example, Gosh 2006). Yet the contagion of the current financial crisis has been as painful as that of the homegrown capital account crisis of 1997 and 1998.

During the height of the crisis in the fourth quarter of 2008, some countries such as the Republic of Korea saw their financial systems pushed to the brink of insolvency despite the fact that they had ample foreign exchange reserves ($260 billion or 24 percent of gross domestic product, GDP, in Korea) at the beginning of the crisis. These economies have recovered from the liquidity crisis rather quickly, but the crisis is far from over, and their turnaround should not be taken as prima facie evidence that further financial reform is not needed.
This paper analyzes the causes and consequences of the contagion of the global financial crisis—in particular, the reserve currency liquidity crisis—with a view to identifying an agenda for reform of domestic, regional, and global regulatory systems from the perspective of East Asia’s emerging economies. One of the major culprits in the current liquidity crisis is the mismatched maturity of foreign assets and liabilities in bank balance sheets. As Brunnermeier and others (2009, 38) point out, maturity mismatches are not confined to emerging economies. In fact, the mismatch “has been the main source of instability in this and previous financial crises.” In emerging economies the maturity mismatch can have more serious consequences, as it invariably entails a currency mismatch and is exacerbated by procyclicality of capital flows.

The vulnerability of emerging economies to the twin crises is not new. What is new is that some of the existing evidence suggests that, contrary to conventional wisdom, a large reserve holding, a more flexible exchange rate system, current account surpluses, and regulatory restrictions limiting the incidence of maturity and currency mismatches in bank balance sheets have not provided emerging economies with a sturdy shield against external shocks such as the U.S. subprime crisis.

This paper argues that the two balance sheet mismatches arise mostly from normal banking operations and hence are largely unavoidable and that domestic regulations alone are not going to be the most effective way of preventing them, more so when cross-border financial transactions are fully liberalized. What is needed are new global institutions, new financial supervision and regulation, and a global lender of last resort.

These new institutions need to be complemented by the prudential regulation of capital movements in emerging economies. Past experiences with global regulatory reform cast doubt as to whether the new system could be established in the first place and, if created, whether it could be effective. If it could not be created, then second-best solutions will have to be found at the regional level to help emerging economies prevent or manage future crises better.

The paper examines the impact of the global economic crisis on East Asia, provides data on the degree of the two mismatches in several of East Asia’s emerging economies, discusses why the regulatory restrictions on the two mismatches have not been enough to mitigate the impact of the crisis, and explores the areas where further reforms are called for in the national, regional, and global financial systems. A final section concludes.

Impact of the Global Economic Crisis on East Asia’s Financial Markets and Institutions

Amid the deepening global economic crisis, East Asia appeared to have been better sheltered from the global crisis than other regions as late as November 2008. Many forecasts including the November update of the World Economic Outlook (IMF 2008) were upbeat about East Asia’s prospects, suggesting that, as a whole, the region would deliver robust growth, while the United States and the euro area
would struggle with a contraction of their economies in 2009. These optimistic forecasts, however, turned out to be premature and had to be revised downward to project a dismal outlook for 2009.

For more than six months after the collapse of Lehman Brothers in September 2008, the worsening of the global financial crisis led global investors and lenders to deleverage and seek a safe haven in U.S. treasuries, thereby causing equity flows and bond issuance to plunge in global capital markets and curtailing the availability of short-term finance in emerging Asia. As a result, stock prices nosedived throughout the region (see figure 1), and exchange rates, except those pegged to the U.S. dollar, experienced a sharp depreciation against the major currencies and exhibited a higher degree of volatility than before (see figure 2). The sovereign spreads widened, and the quality of their financial liabilities denominated in U.S. dollars measured by the credit-default swap (CDS) spreads also deteriorated before starting to improve in February 2009 (see figures 3 and 4). In this bleak landscape of crisis, banks and other financial institutions pulled back from their lending operations by recalling existing loans instead of granting new ones, as the availability of both local and foreign currency liquidity evaporated, future economic prospects looked dim, and losses piled up.

The earlier optimistic forecast rested on some of the unique features of East Asia. Unlike the 1997–98 Asian crisis, the current crisis was an external shock, with the epicenter located in the United States. East Asia’s economic fundamentals were reck-

**FIGURE 1.**
Movements in Stock Price Indexes in East Asia, 2007–09

```
```
oned to be sound. Compared to those in Europe, the region’s financial institutions were healthier and more profitable because they held fewer U.S. toxic assets. Most countries in the region were also running current account surpluses and accumulating large amounts of foreign exchange reserves. Capping it all, the decoupling of East Asia from the cyclical movements of the rest of the world seemed to have taken root (Anderson 2007; World Bank 2007). Yet by January 2009 it looked as though the sky was falling, and much of East Asia slid into a deeper recession.

To the surprise of many, however, the gloomy outlook did not last long. Since April 2009, emerging East Asia has sprung back, returning to rapid growth and providing hope that the region will lead the global economy out of the crisis.  Debate on the decoupling of East Asia from the consumers of the United States and Europe has been rekindled, as emerging Asia staged an impressive recovery from the crisis in 2009 (IMF 2010).

The severity of the impact of the crisis has varied from country to country. China does not appear to have been affected by the global recession: it posted 9.1 percent growth in 2009. The five original members of the Association of South East Asian Nations (ASEAN) have held up much better than neighboring countries, registering

![Figure 2: Exchange Rates against the U.S. Dollar in East Asian Economies](http://aric.adb.org/macro_indicators.php)
FIGURE 3.
CDS Premium in East Asia, 2008–15

Source: Korea Center for International Finance.

FIGURE 4.
Sovereign Spreads in East Asia, 2008: Foreign Currency–Denominated Sovereign Bond Spreads (versus U.S. Treasury Note)
1.7 percent in the same year. Asian newly industrialized economies—Hong Kong, China; Korea; Singapore; and Taiwan, China—have managed a strong rebound.\(^6\)

The impact of the global economic crisis has been transmitted to East Asia through trade and financial market channels. On the trade channel, the sharp decline in exports to the markets of the United States and European Union has been the main culprit in the current economic slowdown. On the financial channel, the region’s growing and diversified financial ties with the global financial system have allowed turbulence in a financial center country to be passed quickly to the region, destabilizing East Asia’s financial markets and institutions.

As a whole, East Asia is a net lender to the rest of the world, but this position has hardly assured East Asia of its ability to ward off an external financial crisis. The reasons are twofold. One is that East Asia has been importing safe financial assets to augment its foreign exchange reserves while exporting risky ones (Crockett 2000). Given this asymmetry, capital will flow out of the region as a result of global deleveraging and the flight to quality of both foreign and domestic investors in search of less risky assets. East Asian governments have amassed large amounts of foreign reserves, the bulk of which are held in safe assets such as U.S. treasuries. Since the 1997–98 crisis, East Asian banks and financial institutions have become more risk averse in managing their assets and liabilities. They also have been subject to more stringent supervision and regulations designed to safeguard their soundness and safety.

Due in part to these developments, the share of foreign financial instruments in bank portfolios has been relatively small in general; smaller still has been the share of risky structured over-the-counter products, simply because the net return (adjusted for funding cost) on lending to local borrowers has been higher than the return on assets denominated in U.S. dollars or euros. As a result, when the crisis erupted, investors from outside the region began divesting themselves of risky East Asian assets, but East Asian investors—private as well as institutional—have not retreated from global markets simply because they are holding mostly relatively risk-free dollar- and euro-denominated assets. The consequence of this asymmetric diversification has been a steep decline in capital inflows into the region.

Another reason for East Asia’s inability to fend off the crisis stems from its heavy dependence on global financial intermediation conducted mostly in reserve currencies—the U.S. dollar and euro—and dominated by global financial institutions and markets located in the United States and Europe. Under these arrangements, when foreign lenders and investors sell off their East Asian assets and refuse to renew their short-term loans, East Asia’s financial institutions and markets, unlike those of the United States and European Monetary Union members, cannot make up for the shortages of liquidity denominated in reserve currencies. Because of this inability, the squeeze on U.S. dollar liquidity has inflicted a great deal of damage on the wider economy and deepened the recession, bringing about the collapse of foreign currency lending, including trade financing, which has threatened the solvency of many East Asian banks unable to roll over their external loans.

In the face of dwindling reserve currency liquidity, East Asian banks could liquidate some of their holdings of foreign assets, but, since they did not hold many of these assets, liquidation could not cover the loss of liquidity. Furthermore in some
countries—notably Korea and Singapore—a growing share of bank liabilities consisted of short-term loans from foreign banks and wholesale market funding, whereas a large share of their assets, including loans to local borrowers, had longer maturity. This maturity mismatch between foreign assets and liabilities exacerbated the liquidity crunch, exposing the entire financial system to a systemic risk of insolvency and eventually a run on central bank reserves—that is, a currency crisis.7

How Serious Were Maturity and Currency Mismatches in East Asia?

During the 1997 Asian crisis, currency mismatches between foreign currency assets and liabilities in bank balance sheets were the major financial vulnerability of Asian banks that exacerbated, if not triggered, the financial meltdown. Various studies argue that currency mismatches played a central role in the 1997–98 Asian financial crisis (Chang and Velasco 2000; Corsetti, Pesenti, and Roubini 1999; Rodrik and Velasco 1999). Goldstein and Turner (2004) argue that all prominent financial crises in emerging economies in the 1990s and early 2000s share one striking characteristic: a large currency mismatch. Most of these studies find the causes of the currency mismatch in market failures associated with asymmetric information and moral hazard. In the 2008–09 crisis, currency mismatches have been relatively mild compared to the massive deterioration in the run-up to the 1997–98 crisis, although the same cannot be said about maturity mismatches.

Maturity mismatching has been a major cause of the crisis not only in emerging but also in advanced economies. Brunnermeier and others (2009) point out that one of the most critical lessons of the 2008–09 crisis is that maturity mismatch—short-term funding of long-term assets with potentially low market liquidity—is a main source of financial instability. In emerging economies with foreign currency liabilities, maturity mismatches create a more serious systemic risk because they are invariably accompanied by currency mismatches.

Causes of the Two Mismatches

There are three major causes of the twin mismatches in emerging economies. One is the role of banks in the transformation of debt maturity, another is procyclicality in bank lending and borrowing, and the third is relationship banking, in which banks establish long-term relationships with their loan customers.

Debt-Maturity Transformation

All banks, whether they are operating from advanced or emerging economies, are essentially engaged in debt-maturity transformation. Banks earn a substantial share of their profits by borrowing from the short end of the financial market (for example, accepting short-term deposits and issuing certificates of deposit) and lending long (for example, extending mortgages to households and loans to business firms for long-term investment in addition to short-term working capital).8 According to Allen and Gale
(2007, 59), the maturity mismatch “reflects the underlying structure of the economy in which individuals have a preference for liquidity but the most profitable investment opportunities take a long time to pay off. Banks are an efficient way of bridging the gap between the maturity structure embedded in the technology and liquidity preference.”

It would be reasonable to assume that under normal circumstances, an individual bank would have an adequate base of deposits and access to wholesale funding markets to finance its long-term loans and investments in securities. In fact, most banks would make the same assumption because they would suffer a competitive disadvantage otherwise. However, in a crisis situation, when depositors leave banks en mass and liquidity in short-term funding markets dries up suddenly, banks often have no choice but to turn to the central bank—lender of last resort—for rescue financing.

**Procyclicality**

The twin mismatches are often exacerbated by the procyclicality in the lending behavior of banks and other financial institutions (Crockett 2000; Borio 2003; White 2004). When the economy enters an upswing phase of the business cycle, financial institutions expand their lending more than before in the belief that the credit risk of their loans has decreased, with the bulk of funds raised from relatively cheaper wholesale funding markets at home and abroad. During the expansionary phase, lending for the financing of housing and commercial real estate is often the major cause of a boom and a bubble in the real estate market. The credit expansion feeds and is often fed by the real estate market boom. Banks and other financial institutions may realize that their excessive lending may indeed create an asset market boom, sowing the seeds of a bubble, which will eventually burst. It would be in their interests to restrain their lending collectively, but there is no market mechanism capable of bringing about such collective actions among financial institutions.

Eventually the expansionary phase or the boom comes to an end, and the economy cools off. At this point, financial institutions become conscious of the potential increase in the credit risk of their loans and begin to recall the existing loans, while refusing to extend new credit as the price of assets, which are in part held as collateral, begins to fall. For an individual institution, cutting credit exposure is a rational decision, but if all institutions do the same, they end up deepening the contraction.9

Over the business cycle, the procyclicality aggravates the two mismatches and amplifies the boom-bust cycle in the market for bank loans. During the boom period, banks have incentives to rely more on short-term funding, as the yield curve is upward sloping. Banks borrow more in volume and from the short end of both domestic and international money markets. The maturity and currency mismatches deteriorate. During a downturn, they do the opposite. But domestic banks and other financial institutions cannot easily recall their foreign currency loans to local customers, and they find it difficult to roll over their borrowings from foreign lenders. As a result, net capital inflows decline sharply, exacerbating the currency mismatch and possibly provoking a reserve currency liquidity crisis. As shown by Kaminsky, Reinhart, and Vegh (2004), Contessi, De Pace, and Francis (2008), and Cardarelli, Elekdag, and Kose (2009), capital flows in emerging markets tend to be procyclical.
During the downturn, the central bank is expected to loosen monetary policy to prevent credit contraction and allow the foreign exchange rate to depreciate. However, expansionary monetary policy can be counterproductive, as it induces further capital outflows. In most emerging economies, equity flows, which account for a large share of capital flows, are insensitive to changes in the exchange rate, so currency depreciation does not significantly restrain capital outflows (Park 2009).10

Even when macroprudential supervision and other regulations have been in place, financial market deregulation and opening have exacerbated maturity and currency mismatches. Financial liberalization has led to the creation and rapid growth of a large variety of short-term money market instruments. Attracted by their relatively high yields, bank depositors have moved out of banks in increasing numbers and into money markets, thereby eroding the deposit base. With this erosion of the traditional funding base, banks have been forced to rely more on both domestic and international wholesale funding markets.

**Relationship Banking**

Banks are in general relationship lenders. Banks develop close relationships with borrowers over time to facilitate monitoring and screening and to overcome problems of asymmetric information. When they refuse to roll over short- as well as long-term loans, they run the risk of losing their loan customers with good credit. Banks also know that most of their loan customers are so accustomed to the loan rollover that they are not prepared to repay their loans even when they are due, let alone pay them back before they mature. Therefore, until they exhaust all other options, banks do not consider refusing to renew most of their household and business loans. Instead, in order to avoid losing customers with a long-standing relationship and liquidating assets, banks turn to the central bank for liquidity support even at a penalty rate of interest.

If domestic currency liquidity is constrained, the central bank can avert a liquidity crisis by pumping more money into the economy. In contrast, when foreign currency liquidity dries up, banks run into the same difficulty: they are unable to sell foreign assets or recall foreign currency loans to their local customers. But unlike in the case of shortages of domestic currency liquidity, the central bank can meet only a limited amount of the increase in demand for foreign currency liquidity. In this case, banks run up both maturity and currency mismatches at the same time. When faced with a sharp decrease in net capital inflows, some individual banks may be able to avoid a liquidity crisis, but the financial system as a whole cannot.

**Measurement of Maturity and Currency Mismatches**

In the aftermath of the 1997–98 crisis, East Asian economies—in particular those hit by the crisis—made concerted efforts to improve the efficiency and stability of their financial systems. Banks and non-bank financial institutions strengthened risk management, improved governance, and fortified themselves with equity capital more than what was needed to meet the Bank for International Settlements capital ade-
quacy requirements. On the macroeconomic policy front, they embraced more flexibility in managing the exchange rate system. To complement these reform measures, they also amassed large amounts of foreign exchange reserves for self-insurance against future crises. Yet they were hardly immune to liquidity risk when foreign lenders and investors liquidated their investments in the domestic financial assets of or refused to renew their loans to East Asian banks.

This section examines the scale and pervasiveness of currency and maturity mismatches in East Asia’s banking industry, which may help to verify whether regulatory restrictions could relieve these balance sheet conflicts. Unfortunately, it is difficult to obtain the micro-banking data needed to estimate the extent of the two mismatches at the regional level. The variables chosen for the examination are at best crude measures of the extent of the two mismatches.

Table 1 presents aggregate effective currency mismatches (AECMs) of East Asia’s emerging economies estimated by Goldstein and Turner (2004). Recent figures are provided by Philip Turner at the Bank for International Settlements. The estimates show that Indonesia, Korea, the Philippines, and Thailand all blundered by letting the currency mismatch rise beyond a safe level in the run-up to the 1997–98 crisis. Since then, the AECM has gradually declined, while remaining in positive territory in most countries. Korea is an exception: its AECM has declined since 2005, turning negative in 2008 due to a sharp decline in net foreign assets. Reflecting the intensity of the 2008–09 global economic crisis, currency mismatching deteriorated in all of these ASEAN countries in 2008.


<table>
<thead>
<tr>
<th>Year</th>
<th>China</th>
<th>Indonesia</th>
<th>Korea, Rep. of</th>
<th>Malaysia</th>
<th>Philippines</th>
<th>Taiwan, China</th>
<th>Thailand</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>3.4</td>
<td>−9.5</td>
<td>−0.5</td>
<td>3.5</td>
<td>1.6</td>
<td>4.7</td>
<td>−2.7</td>
</tr>
<tr>
<td>1995</td>
<td>3.7</td>
<td>−8.6</td>
<td>−1.3</td>
<td>2.3</td>
<td>0.8</td>
<td>3.7</td>
<td>−7.1</td>
</tr>
<tr>
<td>1996</td>
<td>4.9</td>
<td>−8.1</td>
<td>−5.3</td>
<td>1.5</td>
<td>−2.0</td>
<td>3.6</td>
<td>−10.7</td>
</tr>
<tr>
<td>1997</td>
<td>7.0</td>
<td>−21.6</td>
<td>−11.1</td>
<td>−0.9</td>
<td>−6.7</td>
<td>3.4</td>
<td>16.2</td>
</tr>
<tr>
<td>1998</td>
<td>6.1</td>
<td>−16.8</td>
<td>−3.0</td>
<td>2.3</td>
<td>−7.0</td>
<td>3.8</td>
<td>6.3</td>
</tr>
<tr>
<td>1999</td>
<td>4.6</td>
<td>−7.6</td>
<td>1.8</td>
<td>3.6</td>
<td>−6.9</td>
<td>3.8</td>
<td>−0.2</td>
</tr>
<tr>
<td>2000</td>
<td>3.8</td>
<td>−2.2</td>
<td>2.2</td>
<td>2.0</td>
<td>−11.7</td>
<td>2.6</td>
<td>1.1</td>
</tr>
<tr>
<td>2001</td>
<td>4.0</td>
<td>0.7</td>
<td>2.6</td>
<td>2.1</td>
<td>−14.8</td>
<td>4.7</td>
<td>3.0</td>
</tr>
<tr>
<td>2002</td>
<td>2.5</td>
<td>2.5</td>
<td>2.0</td>
<td>1.0</td>
<td>−14.6</td>
<td>7.2</td>
<td>3.9</td>
</tr>
<tr>
<td>2003</td>
<td>2.4</td>
<td>2.9</td>
<td>2.9</td>
<td>2.3</td>
<td>−18.6</td>
<td>12.0</td>
<td>3.9</td>
</tr>
<tr>
<td>2004</td>
<td>3.1</td>
<td>1.7</td>
<td>3.2</td>
<td>3.3</td>
<td>−16.7</td>
<td>11.8</td>
<td>3.4</td>
</tr>
<tr>
<td>2005</td>
<td>3.5</td>
<td>2.2</td>
<td>2.5</td>
<td>2.9</td>
<td>−14.0</td>
<td>11.0</td>
<td>3.7</td>
</tr>
<tr>
<td>2006</td>
<td>3.7</td>
<td>3.5</td>
<td>1.9</td>
<td>3.8</td>
<td>−5.8</td>
<td>9.1</td>
<td>4.2</td>
</tr>
<tr>
<td>2007</td>
<td>4.6</td>
<td>4.5</td>
<td>1.0</td>
<td>4.3</td>
<td>1.0</td>
<td>9.7</td>
<td>4.3</td>
</tr>
<tr>
<td>2008</td>
<td>—</td>
<td>3.0</td>
<td>−0.6</td>
<td>3.4</td>
<td>0.8</td>
<td>9.2</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Sources: International Monetary Fund, Bank for International Settlements, and national data.
Note: This table updates table 4.5 in Goldstein and Turner (2004). — = not available.
As Goldstein and Turner admit, the AECM is an approximate measure. It does not fully reflect the scope of a liquidity crisis largely because it does not take into account differences in the maturity of foreign assets and liabilities. Even when the AECM is positive, a country can experience a foreign currency liquidity crisis if it is exposed to a large maturity mismatch between foreign currency assets and liabilities. To make it more practical, therefore, the AECM needs to be adjusted for the maturity mismatch.

Unfortunately, much of the micro-banking data needed to construct a measure of mismatch for all East Asian countries are not readily available. In their absence, this paper uses changes in the loan-deposit ratio and short-term foreign liabilities relative to foreign exchange reserves to qualify rather than quantify the extent of maturity mismatch. In general, a rise in the loan-deposit ratio indicates that banks rely more on both domestic and foreign wholesale market funding than on core deposits.

An increase in short-term foreign liabilities relative to foreign exchange reserves is likely to be the result of an increase in banks’ external funding from the short end of global financial markets. Here one could use the volume of total foreign liabilities rather than the level of foreign exchange reserves as a scale variable. This paper chooses the latter because the ratio of short-term foreign liabilities to foreign exchange reserves is also regarded as a measure of an adequate amount of foreign exchange reserves to be held in emerging economies. A large increase in the share of short-term external indebtedness may worsen currency mismatching but may not necessarily set off capital outflows and a liquidity crisis if a country holds a large amount of foreign exchange reserves.

As shown in table 2, loan-deposit ratios have been stable and have remained well below 100 percent in most countries, suggesting that as a whole, East Asian banks have had a sufficient base of deposits to meet the local demand for loans. Exceptions are Korea and Thailand, where the ratios climbed up to 127 and 105 percent, respec-

### TABLE 2. Loan-Deposit Ratio in Select Asian Economies, 2000–08

<table>
<thead>
<tr>
<th>Year</th>
<th>China</th>
<th>Indonesia</th>
<th>Korea, Rep. of</th>
<th>Malaysia</th>
<th>Philippines</th>
<th>Singapore</th>
<th>Taiwan, China</th>
<th>Thailand</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>80.3</td>
<td>37.3</td>
<td>87.1</td>
<td>84.3</td>
<td>76.1</td>
<td>87.8</td>
<td>83.0</td>
<td>93.3</td>
</tr>
<tr>
<td>2001</td>
<td>78.2</td>
<td>38.0</td>
<td>88.7</td>
<td>85.9</td>
<td>74.0</td>
<td>87.3</td>
<td>75.1</td>
<td>84.1</td>
</tr>
<tr>
<td>2002</td>
<td>76.8</td>
<td>43.2</td>
<td>97.1</td>
<td>84.9</td>
<td>70.8</td>
<td>87.8</td>
<td>72.3</td>
<td>88.1</td>
</tr>
<tr>
<td>2003</td>
<td>76.4</td>
<td>48.5</td>
<td>101.2</td>
<td>80.9</td>
<td>72.3</td>
<td>86.3</td>
<td>73.4</td>
<td>85.9</td>
</tr>
<tr>
<td>2004</td>
<td>73.7</td>
<td>57.4</td>
<td>101.5</td>
<td>78.6</td>
<td>65.7</td>
<td>84.2</td>
<td>77.6</td>
<td>89.2</td>
</tr>
<tr>
<td>2005</td>
<td>67.8</td>
<td>60.8</td>
<td>103.6</td>
<td>77.5</td>
<td>63.5</td>
<td>79.5</td>
<td>79.8</td>
<td>88.6</td>
</tr>
<tr>
<td>2006</td>
<td>67.2</td>
<td>60.6</td>
<td>112.3</td>
<td>70.8</td>
<td>59.0</td>
<td>69.2</td>
<td>79.2</td>
<td>86.9</td>
</tr>
<tr>
<td>2007</td>
<td>67.2</td>
<td>65.1</td>
<td>128.7</td>
<td>72.2</td>
<td>58.5</td>
<td>71.2</td>
<td>80.4</td>
<td>91.4</td>
</tr>
<tr>
<td>2008</td>
<td>65.1</td>
<td>73.2</td>
<td>127.4</td>
<td>73.5</td>
<td>58.4</td>
<td>75.6</td>
<td>75.6</td>
<td>105.0</td>
</tr>
</tbody>
</table>

tively, in 2008. And Indonesia saw a large hike in 2008, with the ratio rising to 73 percent, up from 65 percent the preceding year.

On the external liability side, short-term foreign indebtedness as a proportion of foreign exchange reserves in some of the East Asian countries where data are available has been well below the level prescribed by the Greenspan-Guidotti-Fischer (GGF) rule, which is to hold an amount of reserves equal to the country’s short-term foreign currency liabilities, a level sufficient to fend off a speculative attack (see tables 3 and 4). The definition of short-term foreign liabilities varies from data source to source. Here two sets of data from two different sources are presented for comparison. The ratios have risen substantially in Indonesia, Korea, and Singapore but have remained below 100 percent.12

The two ratios, together with changes in the AECM, suggest that except for Korea, Indonesia, and possibly Singapore, other countries have so far remained outside the danger zone of a currency crisis. But the 2008–09 global financial crisis is far from over and could flare up again, depending on the effectiveness of international efforts to stimulate the global economy.

### TABLE 3. Short-Term Foreign Liabilities as a Percentage of Foreign Exchange Reserves in Select Asian Economies, 2005–08

<table>
<thead>
<tr>
<th>Year</th>
<th>China</th>
<th>Indonesia</th>
<th>Korea, Rep. of</th>
<th>Malaysia</th>
<th>Philippines</th>
<th>Singapore</th>
<th>Taiwan, China</th>
<th>Thailand</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>14</td>
<td>85</td>
<td>31</td>
<td>17</td>
<td>56</td>
<td>142</td>
<td>29</td>
<td>31</td>
</tr>
<tr>
<td>2006</td>
<td>13</td>
<td>77</td>
<td>48</td>
<td>14</td>
<td>39</td>
<td>162</td>
<td>30</td>
<td>27</td>
</tr>
<tr>
<td>2007</td>
<td>10</td>
<td>41</td>
<td>61</td>
<td>16</td>
<td>38</td>
<td>170</td>
<td>37</td>
<td>25</td>
</tr>
<tr>
<td>2008</td>
<td>9</td>
<td>73</td>
<td>75</td>
<td>19</td>
<td>33</td>
<td>183</td>
<td>36</td>
<td>22</td>
</tr>
</tbody>
</table>

Sources: Bloomberg and Fitch.

### TABLE 4. Short-Term External Liabilities as a Percentage of Foreign Exchange Reserves in Select Asian Economies, 2000–08

<table>
<thead>
<tr>
<th>Year</th>
<th>China</th>
<th>Indonesia</th>
<th>Korea, Rep. of</th>
<th>Malaysia</th>
<th>Philippines</th>
<th>Singapore</th>
<th>Taiwan, China</th>
<th>Thailand</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>12.9</td>
<td>74.3</td>
<td>43.5</td>
<td>29.1</td>
<td>63.4</td>
<td>84.1</td>
<td>12.3</td>
<td>39.9</td>
</tr>
<tr>
<td>2001</td>
<td>10.4</td>
<td>63.8</td>
<td>40.2</td>
<td>31.7</td>
<td>69.6</td>
<td>89.9</td>
<td>9.7</td>
<td>36.4</td>
</tr>
<tr>
<td>2002</td>
<td>7.3</td>
<td>43.9</td>
<td>36.9</td>
<td>30.2</td>
<td>57.0</td>
<td>71.7</td>
<td>8.6</td>
<td>25.8</td>
</tr>
<tr>
<td>2003</td>
<td>7.7</td>
<td>38.5</td>
<td>34.6</td>
<td>24.6</td>
<td>83.8</td>
<td>58.8</td>
<td>12.1</td>
<td>24.8</td>
</tr>
<tr>
<td>2004</td>
<td>7.0</td>
<td>47.6</td>
<td>28.6</td>
<td>26.4</td>
<td>74.9</td>
<td>60.8</td>
<td>16.1</td>
<td>22.8</td>
</tr>
<tr>
<td>2005</td>
<td>7.7</td>
<td>56.4</td>
<td>29.9</td>
<td>24.9</td>
<td>80.2</td>
<td>71.6</td>
<td>15.0</td>
<td>25.3</td>
</tr>
<tr>
<td>2006</td>
<td>7.5</td>
<td>53.8</td>
<td>42.0</td>
<td>27.6</td>
<td>52.1</td>
<td>75.5</td>
<td>12.5</td>
<td>22.2</td>
</tr>
<tr>
<td>2007</td>
<td>8.0</td>
<td>53.2</td>
<td>56.5</td>
<td>22.4</td>
<td>43.4</td>
<td>79.8</td>
<td>13.7</td>
<td>11.6</td>
</tr>
<tr>
<td>2008</td>
<td>5.4</td>
<td>57.1</td>
<td>58.1</td>
<td>24.0</td>
<td>24.5</td>
<td>67.8</td>
<td>8.5</td>
<td>9.9</td>
</tr>
</tbody>
</table>

Can Maturity and Currency Mismatches Be Mitigated by Regulation?

Banks are drawn into currency mismatches because they finance some of their local currency loans with foreign currency funds; even when they relend foreign currency funds to their local customers, they often commit a currency mismatch because local borrowers include not only exporters with foreign currency cash flows but also importers without such flows, and those borrowers in foreign currency are not prepared for an unexpected recall or denial of the rollover of their loans. The maturity mismatch makes banks vulnerable to a sudden change in the demand for liquidity. However, this is not the end of the story. When combined with a currency mismatch, maturity mismatch can easily cause the local currency to depreciate. The weakening of the currency then worsens currency mismatching further and could trigger a currency crisis.

Private Precautionary Measures

Banks in East Asia, in particular those susceptible to the twin mismatches, are required to take precautionary measures to avoid liquidity risks. In general, individual banks have four options to consider when preparing for and managing foreign currency liquidity shortages: (a) holding reserves in terms of liquid foreign assets, (b) securing contingent lines of credit from foreign banks, (c) securitizing and marketing in global financial markets their loans to local customers, and (d) obtaining foreign currency loans from the central bank. Will these precautionary measures be reliable and effective in guarding against a reserve currency liquidity crisis? Will they help to prevent a systemic risk such as a run on central bank reserves?

On the first option, the share of foreign securities held by banks in East Asia's emerging economies is relatively small. This is because the net return on investing in these assets is likely to be low, as their funding costs are higher than those of their competitors from advanced economies. They will find it more attractive to extend foreign currency loans to domestic customers instead of investing in short-term foreign financial instruments. Furthermore, liquidation of these assets would incur heavy losses since they are likely to be sold at fire-sale prices if they can be sold at all. This option has systemic implications in that if all banks try to sell their foreign assets, liquidity shortages will be worse for the banking sector as a whole.

As for the second option, the fee for contingent credits to be drawn in case of a financial crisis could also be high, as it is likely to reflect the solvency risk to which banks from emerging economies are exposed. More important, contingent lines of credit may spare individual banks a liquidity crisis, but not the entire banking sector. When foreign banks conduct their lending operations within a preset country exposure, they are likely to recall other loans or refuse to extend new loans to compensate for the drawdown of contingent credits.

The third option is also costly. In theory, some of the local as well as the foreign currency loans to local borrowers held in bank portfolios could be securitized and insured by mono-line or multi-line insurers via the CDS market. It is not clear
whether banks can remain competitive if they have to bear a high CDS spread for the securitization. In most emerging economies, securitization of loans and other assets is a new financial innovation in which banks have little expertise. Even if some of their foreign currency loans to domestic borrowers can be securitized, market prices of these derivative products may not be high enough to cover the funding cost when credit and currency risks are priced properly into the values of these instruments.

For example, Korea’s largest commercial bank, Kookmin Bank, raised $1 billion by the sale of covered bonds in May 2009. Its five-year, 7.25 percent notes were priced to yield 500 basis points more than the mid-swap rate. By then, Korea had largely overcome the liquidity crisis, yet it still had to pay a high cost of borrowing. Even when some of the banks are able to issue securitized products, the systemic risk for the economy as a whole cannot be shifted to foreign investors and lending institutions. This is because foreign holders of these securitized products are likely to dump them, causing a free fall of their prices. The depressed prices will be taken as an indication that the crisis in the country where the products were issued is deeper than expected. This expectation will then provoke further outflows of capital.

Finally, on the last option, the central bank does not stand ready to rescue banks beleaguered by liquidity shortages even when it holds large amounts of foreign exchange reserves simply because it cannot assume the role of a lender of last resort for reserve currency liquidity.

If none of these precautionary measures is available, banks faced with a liquidity crisis will herd into the local foreign exchange market to buy dollars and euros, thereby weakening the local currency. Depending on the expectations of future currency movements, the initial depreciation could precipitate currency speculation, touching off a run on the central bank’s foreign exchange reserve. This could happen if foreign investors head to the exit all at once or if foreign banks suddenly stop rolling over their short-term loans.

### Regulatory Restrictions

A strict regulation designed to prevent currency mismatches would dictate that bank lending and debt contracts be made in the currencies in which deposits are denominated and in which customers earn revenues. In an extreme case, loans to local customers who earn revenues in a local currency should be excluded from banks’ foreign currency lending. Would such a regulation be desirable or, which is more important, enforceable?

According to Goldstein and Turner (2004), regulatory restrictions could be a practical means of mitigating currency mismatches. They propose restrictions such as imposing limits on net foreign exchange positions, foreign exchange liabilities, and bank holdings of foreign currency–denominated securities. They also recommend introducing more restrictive rules for managing liquidity risks and a higher reserve requirement on foreign currency deposits.
More specifically, these regulatory restrictions may include (a) linking the class of assets for which short-term funding is secured to the maturity of the funding, for example, by allowing banks to hold only short-term safe and liquid assets for short-term funding, and (b) imposing a capital charge on financial institutions with funding liquidity risks stemming from the two mismatches (Brunnermeier and others 2009). In a crisis situation, these prudential regulations are likely to be ineffective, as the Korean experience shows.

Is the regulation enforceable? In order to alleviate the twin mismatch problems, Korea’s Financial Supervisory Service (FSS) imposes a foreign currency liquidity regulation in which banks are required to relend in foreign currencies to local borrowers for a minimum of 85 percent of their foreign currency funds maturing within three months (15 percent for domestic currency loans). The maturity of the local foreign currency loans must also be less than three months. In order to prevent the currency mismatch, the FSS also enforces another liquidity restriction in which banks are required to keep the ratio of net short-term foreign currency assets maturing within seven days to total foreign-currency assets at a positive level and the ratio of foreign currency assets maturing in less than 30 days to total foreign currency assets at minus 10 percent.

Banks do comply with these regulatory measures on their balance sheets, but not in reality. Indeed, if these prudential measures had been observed to the letter, Korea should have avoided the run on central bank reserves during the fourth quarter of 2008. But it could not. When global short-term money markets froze, it ran into a serious liquidity squeeze. It appears that banks’ compliance did not prevent, or even moderate, the pervasiveness of the two balance sheet mismatches, largely because the banks kept on renewing their domestic and foreign currency loans regardless of their maturities, with the expectation that they would have continuing access to global wholesale funding markets.

This lax attitude to compliance does not necessarily reflect a serious moral hazard on the part of Korean banks, because, if past experience with the 1997–98 financial crisis is any guide, they know that the government could not come to their aid in a crisis caused by the drought of reserve currency liquidity. Instead, it may have more to do with relationship banking and reflect the fact that compliance means the loss of their competitiveness vis-à-vis their foreign competitors in global financial intermediation.

Banks in reserve currency countries are not subject to these and other restrictions recommended by Goldstein and Turner, because these central banks can always print more money to thwart an impending liquidity crisis caused by a rise in maturity mismatch. Therefore, if carried out too tightly, the regulations in emerging economies could be counterproductive, as they run the risk of limiting the ability of even well-managed banks to compete and to participate in international financial intermediation.

If the regulatory restrictions prove to be ineffective, governments of emerging economies may invoke more direct measures such as providing government guarantees on foreign loans and imposing capital controls. On the government guarantee, a recent Korean experience is instructive. On October 12, 2008, when Korean banks were not able to renew their short-term external loans, the Korean government sought
to restore investor confidence by issuing sovereign guarantees on up to $100 billion of new foreign loans maturing before the end of June 2009. Similar guarantees had failed to allay fears of financial meltdown at the beginning of the Asian crisis in 1997, and they failed again. As they had done in 1997, the markets reacted with indifference.

When Korea secured a swap line amounting to $30 billion from the U.S. Federal Reserve on October 30, 2008, the foreign exchange market settled down somewhat, but not for long. The foreign exchange rate shot up to W 1,509 per dollar three weeks after the swap was announced, which was apparently not enough to remove uncertainties surrounding Korea’s ability to service its foreign debt. On December 13, Korea also managed to arrange won–local currency swaps with the central banks of both China and Japan, each amounting to an equivalent of $30 billion. These measures did not help, but when it became clear that the U.S. Federal Reserve would renew the swap agreement, foreign investors’ confidence in the Korean economy improved, and the foreign exchange market stabilized toward the end of the first quarter of 2009. This is not the place to discuss the benefits and costs of capital controls, but national policy authorities will not stand idly by.

The twin mismatches render these economies highly susceptible to a reserve currency liquidity crisis. In contrast, most advanced countries, in particular reserve currency countries—the United States and members of the European Monetary Union—are largely free from currency mismatches and hence spared the systemic risk of a currency crisis, although they could still suffer a liquidity crisis. This reserve or key currency privilege puts banks in advanced economies at a competitive advantage in international financial intermediation. From the emerging economies’ point of view, this bias, which is equivalent to a non-reserve currency premium for external borrowing to emerging economies, raises the question of unfairness—not to mention the rationale for integrating into the global financial system—to the extent that it creates an uneven playing field in international intermediation business.

Financial Reform: Where and How?

The financial history of the global economy documents a long list of financial crises that have erupted in both advanced and developing countries over the last 400 years (Allen and Gale 2007). This history leaves little doubt that bubbles, excessive exuberance, and crashes are intrinsic features of the market-oriented Western financial system. The current financial crisis will not be the last one the global economy will ever encounter.

Over the past two decades, more and more emerging economies have moved to liberalize their financial markets and integrate their financial intermediation industries into the global financial system, thereby broadening and deepening financial globalization. Whatever its benefits may be, integration with the global financial system has been accompanied by financial instability, as emerging economies are finding it increasingly difficult to deflect adverse external shocks emanating from speculative activities in international financial markets even when their exchange rates are freely floating.
For the efficiency of the global financial system, market globalization needs to be backed by the globalization or even harmonization of financial supervision and regulation and the creation of a global lender of last resort. A global regulatory institution could help to uncover financial market risks that require policy responses on the part of national policy and regulatory authorities. It could enforce uniform global regulatory standards to minimize regulatory arbitrage arising from the existence of different regulatory systems in different countries. It could mobilize coordinated global efforts to prevent a crisis from spreading in a country or region to other parts of the world. Runs on banks and central bank reserves in emerging-market economies caused by excessive speculation, panic, and herding among market participants can only be averted or at least better managed if there were a global lender of last resort.

Despite the repeated calls for more effective and coordinated global responses to financial crises, past experiences with global financial reform cast doubt as to whether these global institutions could be established in the first place and, if they were created, whether they would be effective in improving global financial governance. Creating a globally integrated regulatory system is highly unlikely because there is no one-size-fits-all regulatory institution appropriate to advanced, emerging, and developing economies alike and because financial regulation and supervision come naturally under the purview of national governments insofar as they are required to pay for the cost of crisis resolution. Ten years after creating the euro, the European Monetary Union member countries have yet to unify their national regulatory systems.18

If creating a global regulatory institution is impractical, what is the second-best solution? According to the G-20 (Financial Stability Board 2009), it is to raise global financial standards in a consistent way that ensures a level playing field and avoids fragmentation of markets, protectionism, and regulatory arbitrage. To this end, the Financial Stability Board (FSB) expects to put in place a framework to strengthen adherence to international regulatory and prudential standards by the end of 2010. It is also expected that the FSB members will implement international financial standards and agree to subject themselves to periodic peer reviews.

But it is again unrealistic, and would be futile to attempt, to establish uniform regulatory standards acceptable to all countries regardless of their stage of financial development simply because many emerging economies may not yet have the institutional capacity to adopt and enforce global standards in line with country-specific needs. Recognizing this gap, the G-20 proposes that advanced economies, the International Monetary Fund, and other international organizations provide capacity-building programs for emerging-market economies and developing countries. Since capacity building is likely to take an extended period of time, the G-20 implicitly would allow emerging and developing economies to keep their current standards, even if they are not consistent with the new global standards to be introduced.

Will the new G-20 initiative bear fruits? It is too early to tell, but the disappointment with similar initiatives in the past does not bode well for its future success. After the Asian crisis of 1997–98, many international forums were established to create global standards on accounting, disclosure, risk management, and corporate
governance at banks and other financial institutions and also to set rules for financial supervision and regulation. Over time the memories of the crisis faded, and so did the initial thrust of these global initiatives.

In the preceding section it is argued that prevention of excessive maturity and currency mismatches is a top priority for regulatory reform in emerging economies. It requires keeping in check procyclicality of capital flows and cooperation with and support of financial policy and regulatory authorities of advanced economies. The G-20 supports higher capital requirements for risky products and off-balance-sheet activities and countercyclical capital buffers for moderation of procyclicality. Holding the buffers is similar to dynamic provisioning, which requires banks to build up capital, put a brake on the expansion of bank credit during an expansionary phase of the business cycle, and reverse the process when the economy slows down. It also supports the introduction of a leverage ratio as a supplementary measure to the Basel II risk-based framework. To ensure comparability, the details of the leverage ratio will have to be harmonized internationally, fully adjusting for differences in accounting.

Most emerging economies would, in principle, be in broad agreement with these G-20 proposals, but may not find them to be acceptable because they do not address the problem of the non-reserve currency premium they have to bear. Given the premium, the market will demand higher capital requirements and a higher leverage ratio than they do in advanced economies. For example, to alleviate the procyclicality, emerging economies will need to build larger reserves than advanced countries in an upturn of the business cycle, thereby undermining competitiveness of their financial institutions.

As for multilateral cooperation, emerging economies do not expect advanced countries to require their financial institutions to accept liabilities from abroad only if they are denominated in emerging economies’ currencies. But they would benefit from the creation of a multilateral cooperative arrangement for supervision and regulation of financial institutions, which could facilitate exchanges of information between the concerned regulatory authorities from advanced and emerging economies. The cooperative arrangement could be structured and managed in a way that could facilitate and assist the joint supervision and inspection of both lenders from advanced countries active in emerging economies and systemically important borrowers from emerging economies.

For example, when an emerging economy finds a large increase in maturity mismatching of foreign assets and liabilities in its financial institutions, there is the danger that domestic borrowers may commit excessive leveraging and foreign lenders may take undue risk. Under these circumstances, the emerging economy could turn to the new cooperative arrangement for joint supervision of both domestic borrowers and foreign lenders with the regulatory authorities of the countries where the external lenders operate. Such joint supervision may help to provide the emerging economy with the information it needs to take precautionary measures to moderate maturity and currency mismatches by stabilizing interregional capital flows.

Although the proposal is a modest one, it is highly uncertain whether such an institution will ever come into existence. In view of this uncertainty, it would be in the
interests of East Asia’s emerging economies to participate actively and establish a united front in voicing their concerns and priorities in various international forums entrusted with the reform of the global regulatory system, including the G-20, the Financial Stability Forum, and the Basel Committee on Banking Supervision.

**Regional Cooperative Arrangements**

If the prospects for reform of the global regulatory system are not promising, are there any regional arrangements for regulatory cooperation that may help to safeguard East Asia’s emerging economies against turbulences of international finance? There is an East Asian regional cooperative arrangement known as ASEAN+3. Although it has a relatively short history, it has made considerable progress in strengthening regional cooperation by creating a regional liquidity support system under the Chiang Mai Initiative Multilateralization (CMIM).

At present, the CMIM does not play any role in coordinating financial supervision and regulation at the regional level, but it could, once it is armed with its own surveillance functions, which are planned to be added in 2010. ASEAN+3 could then establish a college of the region’s supervisory authorities as an integral component of the CMIM. A regional college of supervisors would offer an effective way of improving communications among the supervisors of those systematically important multinational financial services conglomerates operating out of East Asia. To support the college in this role, the existing ASEAN+3 early-warning system could be improved to alert regulators and financial institutions to emerging risks, including those from macroeconomic trends and volatile capital flows in both the regional and global economies.

**Regulatory Reform at the National Level**

It is obviously too early to conjecture the features of the new global regulatory system that may emerge from the ongoing discussions at the G-20 and other international financial forums. At this stage, therefore, it is premature and risky for East Asia’s emerging economies to launch any major institutional and structural reform of the financial system independently of the financial reforms under way in the United States, Europe, and the G-20. They would be better advised to wait until the contours of new global financial regulations are clear. This does not mean that they can remain on the sidelines of the global regulatory reform efforts. Whatever global system is constructed, its effectiveness will be enhanced if it is complemented by broadening the scope and developing new instruments of domestic macroprudential supervision to bring under control the procyclicality of bank lending and capital flows.

Strengthening macroprudential supervision and expanding regional and global regulatory cooperation will help, but will not be enough to prevent future liquidity crises unless complemented by liquidity support from reserve currency countries. The U.S. dollar is a de facto global currency, and the euro has emerged as a distant-second reserve currency. As the providers of global mediums of exchange and stores
of value, the reserve currency countries need to bear responsibility for controlling and stabilizing the global supply of liquidity. In assuming their global role, the U.S. Federal Reserve (Fed) and the European Central Bank may consider institutionalizing a global currency swap network to be activated in a crisis situation. At present, the central banks of Canada, Japan, Switzerland, and the United Kingdom together with the European Central Bank have unlimited dollar swap lines with the Fed. The network membership could be enlarged to include, in addition to the current members, many of the emerging economies active in international financial markets. These emerging-market countries would then have access to swap lines that could provide short-term U.S. dollar or euro liquidity if needed to prevent a liquidity crisis.21

The severity and contagiousness of the current financial crisis have raised concerns that some of East Asia’s emerging economies may decide to retreat from or at least reduce the speed and scope of financial market opening and integration into the global financial network. Such a decision would be premature and may not serve their interests in the long run. In contrast, other emerging economies have taken steps to liberalize and open their financial markets to facilitate internationalization of their currencies, with the expectation of enlarging the ability to borrow in their own currencies. Although the extent to which internationalization has relieved liquidity shortages in those countries with fully convertible currencies is unclear, it is a new avenue worthy of further research for some of the East Asia’s emerging economies.22

Concluding Remarks

The eruption of a crisis in the subprime mortgage market in the United States in 2007 has passed through to other advanced and emerging economies, causing a financial meltdown in advanced economies and subsequently setting off a severe recession throughout the global economy. Most of East Asia’s emerging economies have been able to ride out the crisis relatively well. But it has been a painful reminder that strong economic fundamentals and a rational macroeconomic policy framework alone may not spare them from liquidity and currency crises precipitated by external shocks such as the U.S. subprime crisis, more so when their financial markets are integrated with the markets of the global financial system.

In order to avoid or mitigate maturity and currency mismatches, which have been some of the main causes of financial vulnerabilities, East Asia’s emerging economies have imposed several micro- and macroprudential regulations on banks’ asset-liability management. At the same time they have moved toward a more flexible exchange rate system, internationalizing their currencies, and accumulating large foreign exchange reserves. Nonetheless, these reform efforts and other precautionary measures may not have been enough to prevent contagion of external financial crises, regardless of the location of their epicenter. When foreign investors and lenders suddenly and unexpectedly withdraw their loans and investments, emerging economies are unable to cover shortages of reserve currency liquidity, which
often trigger a currency crisis, especially when banks are exposed to high degrees of currency and maturity mismatches.

To be sure, maturity mismatches do not plague only emerging economies: they have been as pervasive in the banking industries of advanced countries as they have in other economies. In fact, the mismatching of maturity has been one of the major causes of the financial collapse in advanced countries in 2008. Unlike in emerging economies, however, maturity mismatch in advanced countries is not accompanied by a currency mismatch and hence does not provoke a currency crisis even when global liquidity evaporates. This is because the currencies of the United States and the members of the European Monetary system—the U.S. dollar and the euro—are international media of exchange held as foreign exchange reserves, and other advanced economies can borrow in their own currencies, which are internationalized, from international financial markets. This privilege constitutes a reserve currency benefit that puts advanced economies at a competitive advantage in international financial intermediation. In the absence of this advantage, emerging economies have to pay a premium for their external borrowing.

This disadvantage means there is no level playing field for banks from emerging economies that are engaged in international financial intermediation. This premium is exacerbated when restrictions on foreign currency borrowing and lending or requirements for hedging currency risk are imposed on their financial institutions because these regulations are then translated into a higher cost of funding from international financial markets. From the perspectives of emerging economies, unless institutional arrangements for global liquidity assistance—such as a global liquidity safety net, which is on the G-20 reform agenda—are established to enable them to prevent or better manage liquidity and currency crises, they may have to reexamine the costs and benefits of integrating into the global financial system.

Notes

1. Work related to the regulatory reform is being carried out by many private and public institutions including the G-20, the European Commission, the Financial Stability Board, the International Monetary Fund, the Basel Committee on Banking Supervision, the International Organization of Securities Commissions, and the International Association of Insurance Supervisors. See, for example, FSA (2009), Financial Stability Board (2009), IMF 2009a, 2009b), de Larosière Group (2009), Brunnermeier and others (2009), Bank of England (2009), and Warwick Commission (2009).

2. The most extensive list of reform proposals can be found in Financial Stability Board (2009).

3. Maturity mismatches between assets and liabilities denominated in local currency could also give rise to a liquidity crisis, but such a crisis is within the capabilities of the central bank to resolve.

4. They do not make a distinction between local and foreign currency maturity mismatches.

5. According to a popular journal (Economist, April 15–21), the recovery has been astounding.

6. At this stage of the crisis, the episodes of the ongoing crisis do not provide empirical evidence either supporting or disputing the decoupling phenomenon as well as the notion that East Asia may help to lead a global economic recovery.
7. To be sure, East Asia’s central banks can supply additional liquidity by running down their foreign exchange reserve holdings, which were known to be excessively large before the crisis, but if they do, they run the risk of provoking a run on their reserves if speculators start shorting the local currencies. Except for the yen, few East Asian currencies are internationalized in the sense that they are widely used to settle international transactions in financial assets as well as goods and services. East Asia holds collectively more than $4 trillion in foreign exchange reserves, but it is not clear whether, except for China and Japan, other East Asian economies held enough reserves to ward off currency speculation at the height of the crisis in the fourth quarter of 2008.

8. Brunnermeier and others (2009) argue that there are many caveats to this generalization and that the mismatch is a matter of degree. The incentive to commit maturity mismatch is most pronounced when the yield curve is upward sloping in a boom.

9. This coordination failure may justify intervention on the part of policy authorities. Given the nature of its operations, the supervisory agency may be the authority to assume the market intervention.

10. The supervisory agencies are not specifically entrusted with stability functions; they may not have developed the expertise or culture of macroprudential orientation, while the central bank cannot exercise supervisory control at the level of individual institutions. These institutional constraints could hamper coordination of macroprudential policy, creating the danger that the policy authorities, including the ministry responsible for fiscal and exchange rate policy, will not be able to agree on the seriousness of financial distress once it arises and hence will fail to devise a collective policy response.

11. Goldstein and Turner (2004) were the first to develop a statistical measure of currency mismatch. They define an aggregate effective currency mismatch as follows: \( \text{AECM} = \frac{\text{NFCA}}{\text{XGS}} \times (\frac{\text{FC}}{\text{TD}}) \), where NFCA = net foreign currency assets (+) or liabilities (–) and XGS = exports of goods and services (national income account); when NFCA is negative, MGS = imports of goods and services (national income account); when NFCA is positive, \( \frac{\text{FC}}{\text{TD}} \) = foreign currency share of total debt.

12. A recent Citibank study uses a reserve recovery ratio to measure foreign exchange reserve adequacy, which is defined as the ratio of reserves to the sum of short-term foreign debts by remaining maturity and expected current account surplus or deficit for the next 12 months. Estimates of these ratios for several East Asian countries show that at the end of 2008 Korea had barely enough reserves, with a ratio of 1.1 for 2009, whereas other countries had an ample cushion of reserves, with ratios ranging from 1.6 for Indonesia to 5.4 in Thailand (Huang 2009).

13. Covered loans differ from mortgage-backed securities in that they are secured by property loans or lending to public institutions and are backed by the borrower’s guarantee to make payment.

14. Earlier in April 2009, Hana Bank, Korea’s fourth largest bank, sold $1 billion of three-year government-guaranteed notes priced to yield 490 basis points more than the mid-market swap rate.

15. The mid-market swap rate is the rate at which the discounted future payments of the fixed and floating swap payments net to zero.

16. In addition to the regulatory restrictions, Goldstein and Turner (2004) recommend a managed floating foreign exchange policy to large emerging economies because it produces incentives for banks, non-bank financial institutions, and corporations to hedge currency risk in order to keep currency mismatches under control. However, a recent bout of currency speculation in Korea raises doubts about the extent to which free floating could relieve the burden of currency mismatches at banks and reduce the incidence of a currency crisis when expectations of currency speculators on the future exchange rate tend to be
extrapolative. Holding large amounts of reserves more than what the G-G-F rule prescribes may help to avert a liquidity crisis. When the market sentiment builds up and expectations are firmly held, speculators can hold short positions of any size. In effect, a speculative attack is a run on the reserves of the central bank; the larger the reserves, the bigger the run. In this situation, prices of equities and bonds will continue to fall, and the exchange rate will continue to depreciate until the central bank runs out of its foreign reserves.

17. This is a version of mark to funding proposed by Brunnermeier and others (2009).

18. A similar reasoning would conclude that there will not be a global lender of last resort as long as there are no global taxpayers.

19. The G-20 also sees the need to improve the quality, consistency, and transparency of the Tier 1 capital base in a way that will allow harmonization across jurisdictions and comparisons across institutions to be easily made.

20. The Basel Committee on Banking Supervision has stated that the level of capital in the banking system, both the minimum capital requirement and the buffers above it, will be raised relative to pre-crisis levels to improve resilience to future episodes of stress. This will be done through a combination of measures such as strengthening the risk coverage of the Basel II capital framework, improving the quality of capital, and raising the overall minimum requirement.

21. There is concern that expanding the swap network could create moral hazard problems in emerging economies. But it is difficult to believe that emerging economies would be disposed to laxity in managing macroeconomic policy simply because they have access to the swap lines.


References


Financial Crisis and the Paradox of Underregulation and Overregulation

JOSHUA AIZENMAN

This paper illustrates the paradox of prudential underregulation in an economy that adopts a financial reform that exposes the economy to future financial crises. There is individual uncertainty about the crisis incidence, and the probability of the crisis is updated sequentially applying Bayesian inference. Costly regulation can mitigate the probability of the crisis. The paper identifies conditions where the level of regulation supported by the majority is positive after the reform, but below the socially optimal level. Tranquil time, when the crisis would not take place, reduces the intensity of regulation. If the spell of no crisis is long enough, the level of regulation may drop to zero, despite the fact that the socially optimal level of regulation remains positive, inducing a regulatory cycle. The less informative is the prior regarding the probability of a crisis, the faster is the drop in regulations induced by a no-crisis run of good luck. The challenges facing the regulator are aggravated by asymmetric information, as is the case when the public does not observe the regulator’s effort. Higher regulatory effort, while helping to avoid a crisis, may be misconstrued as a signal that the environment is less risky, reducing the posterior probability of the crisis and eroding the support for costly future regulation. The other side of the regulation paradox is that a crisis resulting in unanticipated high costs may induce overregulation and stagnation, as the parties that would bear the cost of the overregulation are underrepresented in the decision-making process. The paper also outlines a regulatory structure that mitigates these concerns, including requiring information disclosure, increasing the independence of the regulatory agency from the political process, centralizing the regulatory process and increasing its transparency, and adopting global standards of minimum prudential regulations and information disclosure, enforced by the domestic regulator.

Joshua Aizenman is a professor of economics at the University of California, Santa Cruz (UCSC) and a research associate with the National Bureau of Economic Research. He thanks Maury Obstfeld, Dan Friedman, Ken Kletzer, Jong-Wha Lee, Donald Wittman, an anonymous referee, seminar participants at UCSC, and participants at the ABCDE World Bank conference in Seoul for their insightful comments and Yi Sun for dedicated research assistance.

Annual World Bank Conference on Development Economics 2010, Global
© 2011 The International Bank for Reconstruction and Development / The World Bank
“History is the sum total of things that could have been avoided.” Konrad Adenauer

“No less than the tourist, the writer of history profits from maps.” Charles F. Mullett

Following the “lost development decade” of the 1980s, the attitude toward financial integration of developing countries changed dramatically in the 1990s. The growing optimism about the gains from globalization created support for trade liberalization, frequently coupled with financial opening in Latin America and, to a lesser degree, in Asia. Figure 1 plots the broad patterns of financial integration, applying the Chinn-Ito index of de jure financial openness, and figure 2 plots de facto measures of openness (reporting the ratios of net and gross foreign assets to gross domestic product, GDP).\(^1\) While financial liberalization took off in the Organisation for Economic Co-operation and Development (OECD) in the 1980s, that decade was marked by massive sovereign defaults and growing financial isolation of developing countries. In sharp contrast, in the 1990s developing countries joined the OECD trend toward capital openness, a trend that continued globally until mid-2006. Disaggregation reveals the heterogeneity of the patterns of financial openness of developing countries. Latin America moved more assertively to embrace capital openness, while Asia moved timidly, with financial openness reversing around the onset of the East Asian crisis of 1997–98. The attitude toward financial integration of the most populated countries—China and India—remains guarded and skeptical. With a lag, these changes in policies were associated with large increases in the ratio of total foreign assets to GDP and with more modest changes in the ratio of net foreign assets to GDP (although these changes are more pronounced for developing and emerging markets). Econometric evidence suggests the presence of two-way positive “Granger causality” between de facto and de jure measures of financial openness, in line with theoretical models outlining such effects.\(^2\)

FIGURE 1.
De Jure Financial Integration (Chinn–Ito Index), by Level of Development and Region, 1970–2007
a. Emerging, other developing, and OECD countries
The differences in the path toward financial openness taken by Asia and Latin America may reflect the vigorous debate about financial opening. This deliberation is a reincarnation of the earlier literature on immiserizing growth, which identifies conditions under which growth may be welfare reducing in the presence of preexisting distortions. While financial opening increases welfare when the only distortion is to restrict intertemporal trade across countries, financial opening may reduce welfare in the presence of other distortions. An important example of such a distortion is moral hazard, which frequently acts as an implicit subsidy to borrowing and investment. Moral hazard arises when investors believe that taxpayers will bail them out of a bad investment. This bailout may be carried out by the treasury, the central bank,
or international agencies such as the International Monetary Fund (IMF) and the World Bank. In these circumstances, the taxpayer subsidizes the investment. A frequent rationale for the bailout is the “too big to fail” doctrine—the fear that allowing large borrowers to go under will trigger a systemic crisis.³

Less deliberated has been the rapid adoption of financial innovations and financial deregulation in the OECD countries, led by the United States. Frequently, the presumption has been that the superior financial intermediation of the OECD implies that these innovations are welfare improving, with negligible downside risk exposure. The prolonged spell of what was dubbed “the great moderation” further reduced concerns about the downside risk associated with financial intermediation, as reflected by the remarkable decline in the market price of risk. The moderation
also reduced the appetite for regulation, with growing acceptance of Greenspan's seductive “market-stabilizing private regulatory forces” doctrine. Concerns about the inherent principal-agent/moral hazard associated with financial intermediation were swept aside, minimizing the potential role of the regulator. The data also reveal high persistency of the stances of countries toward financial integration: the autoregressive coefficient of AR(1) process fitted to the Chinn-Ito index over time is 0.84 for emerging Asian countries, 0.99 for emerging Latin American countries, 0.93 for non-emerging developing Asian countries, and 1.00 for non-emerging developing Latin American countries. The high co-movement of the Chinn-Ito index across countries and the persistence of policies toward financial openness suggest that the pendulum is swinging toward financial integration. Indeed, the process of globalization increases the interdependency of national capital markets, possibly leading to tighter co-movements of policies and attitudes toward financial openness and other policy dimensions of globalization.

The unfolding global liquidity crisis illustrates vividly the universality of moral hazard and the notion that the “too big to fail” doctrine is shaping government and central bank policies in times of systemic crises. It also serves as a painful reminder of the risk of underregulating domestic and international financial intermediation, exposing the taxpayer to excessive risk taking that in due course would be subsidized by taxpayer-financed bailouts. The magnitude of the global crisis with regard to the exposure to risk, the distribution of toxic assets, and the speed of deleveraging surprised most observers; as of April 2009, the IMF estimates that banks and financial institutions globally will have to dispose of or write down about $4 trillion of toxic assets. The high costs of the crisis suggest that the pendulum of financial integration may shift toward a reversal of financial globalization. Even if we agree with George Santayana that “those who cannot learn from history are doomed to repeat it,” a necessary condition for successful learning is to understand the forces accounting for the past. Simply reversing the policy stances of the past 20 years would backfire, as doing so might overshoot the needed adjustment, inducing other distortions.

This paper seeks to explain the tendency to underregulate in “good times” and the risk associated with overshooting the adjustment needed following a financial crisis. Both underregulation and overregulation may reflect the paradox of financial regulation: the success of the prudential regulator or a prolonged period of economic tranquility lead to complacency, reducing the demand for the regulator's services, inducing underregulation, which leads to a financial calamity. While the identity of economic agents that benefit directly from crisis avoidance is unknown, the cost and the cumbrance of regulations are transparent. Hence crises that have been avoided are imperceptible and underrepresented in the political discourse. The demand for regulation declines during prolonged good times, increasing the ultimate cost of eventual crises. The other side of the regulation paradox is the hazard of overregulating financial intermediation in an environment where the identity of economic agents that benefit directly from financial regulations is known, while the
identity of successful projects and the entrepreneurs who would not be financed due to overregulation is unknown (they are not represented in the political discourse). These considerations suggest the need to strive toward a golden rule of Goldilocks prudential regulations.

This paper considers an economy that adopts a financial reform that exposes the economy to a future crisis. There is uncertainty about the crisis and uncertainty about the probability of future crises. The public’s initial prior is updated sequentially applying a Bayesian inference. Costly regulation determined by a majority rule can mitigate the probability of the crisis. When the majority in the economy is partially exposed to the crisis, and the efficacy of regulation is high enough, the level of regulation following the reform is positive, but below the socially optimal level. Tranquil time, when the crisis would not take place, reduces the intensity of regulation over time. If the spell of no crisis is long enough, the level of regulation may drop to zero, despite the fact that the socially optimal level of regulation remains positive. The less informative is the prior regarding the probability of a crisis, the faster is the drop in regulations induced by a no-crisis run of good luck. The challenges facing the regulator are aggravated by asymmetric information. This would be the case if the probability of a crisis is reduced by the regulator’s effort, but the public does not observe this effort. In these circumstances, higher regulatory effort, helping to avoid a crisis, may be seen as a signal that the environment is less risky, reducing the posterior probability of the crisis. This in turn would reduce the support for costly future regulation. Asymmetric information may put the regulator in an odd position of damned if you do not regulate today, damned tomorrow if you regulate today.

A crisis that creates unanticipated high costs may induce overregulation and stagnation, as the parties that would bear the cost of the overregulation are underrepresented in the decision-making process. The paper closes with an outline of a regulatory structure that mitigates these concerns. Better disclosure of information is needed to allow the regulator to assess in real time the systemic risk triggered by “too big to fail” concerns. Increasing the independence of the regulatory agency from the political process would reduce the tendency to underregulate in good times. Centralizing the regulatory process and increasing its transparency would mitigate the problems associated with asymmetric information. Adopting global standards of minimum prudential regulations and information disclosure that are enforced by the domestic regulator would mitigate the tendency to underregulate in good times.

Financial Reforms, Regulation, and Individual-Specific Uncertainty

This section outlines a minimal model that allows tracing the dynamics of regulation in the aftermath of a financial reform in the presence of individual-specific uncertainty. The key role of individual-specific uncertainty in explaining the status quo bias in the context of trade reform is highlighted by Fernandez and Rodrik (1991), but
seems to be overlooked in explaining the dynamics of regulation in the aftermath of a financial reform. As the experience of the 1990s vividly illustrates, financial reforms increase the economy’s exposure to costly financial crises, events that are associated with asymmetric incidence among agents and sectors. The intertwining of financial institutions operating with limited transparency frequently results in nasty news, when agents find that they are far more exposed to the crisis than they expected. This model embraces the outcome of an interaction between two ingredients. First, over time agents update their assessment of the probability of a crisis, possibly applying a Bayesian inference. Second, regulation is costly. While the costs are shared by all agents, the intensity of regulation is determined by a majority rule. This section studies the dynamics of regulations during a spell of no-crisis episodes in the aftermath of a reform until the crisis happens. It focuses on the case where the majority of agents face individual-specific uncertainty regarding their exposure to incidences of crisis. This is then contrasted to a case in which the intensity of regulation is set in order to maximize the expected utility of the “average agent,” a concept akin to the expected discounted GDP net of regulation costs.

At time 0, a policy reform takes place. The policy reform exposes the economy to the risk of a financial crisis, which will reduce the income of the affected agents by $\tau$. There is uncertainty regarding the probability of such a crisis and its incidence. A recursive structure is assumed in which the intensity of regulation at time $t$ affects the probability of a crisis at time $t + 1$. In the absence of regulation, the perceived probability at time 0 of a crisis occurring at period 1 is $P_{0,1}$. Regulation associated with spending resources $\rho_0$ per agent at time 0 reduces the probability of a crisis from $P_{0,1}$ to $P_{0,1} \cdot Q$, $Q(0) = 1; Q \leq 1$. The efficacy of regulation increases with the spending: $Q = Q(\rho_0)$; $Q' < 0$.

To fix ideas, $Q(\rho_0)$ is assumed to be a logistic function:

$$Q(\rho_0) = \frac{2\exp - \nu \rho_0}{1 + \exp - \nu \rho_0}, \quad (1)$$

where $\nu$ measures the efficacy of regulation. Hence $Q'(\rho_0) = -\nu [Q(\rho_0)/(1 + \exp - \nu \rho_0)]$. Over time, agents update the probability $P_{t, t+1}$ in a Bayesian manner, so that no crisis in period 0 would induce a lower perceived probability of a crisis in period 2, $P_{0,1} \geq P_{0,2}$. The same applies for an arbitrary $t$: no crisis from period 1 to $t$ would imply $P_{t-1, t} \geq P_{t, t+1}$. A detailed example of such inference is provided in the appendix to this chapter.

The economy is populated by a large number of atomistic agents, a fraction $\omega$ of which is fully exposed to the crisis, dubbed $FE$. The remaining population is partially exposed, dubbed $PE$; each agent in the $PE$ group is adversely affected by the crisis with probability $q$. The income in period 0 of all agents is normalized to 1 minus the per capita cost of the regulation, $\rho_0$. The timeline is described in figure 3. Financial reform takes place at the beginning of period 0. Next, the authorities set the regulation intensity, $\rho_0$. At the beginning of period 1, the uncertainty regarding the crisis
and its individual incidences is resolved. With probability $P_{0,1} \cdot Q(\rho_0)$, a crisis will take place in period 1, reducing the income per capita of the affected agents to $1 - \tau$. With probability $1 - P_{0,1} \cdot Q(\rho_0)$, no crisis will take place in period 1.

The key difference between the two groups is the individual’s risk of being exposed to the crisis. In the $FE$ group, all are adversely exposed to the crisis; whereas in the $PE$ group, only a fraction $q < 1$ is affected, and there is individual uncertainty regarding each agent’s exposure. The income profiles of the agents in each group are summarized in table 1.

If no crisis occurs at time 1, the same scenario will repeat itself in period 2, updating the time indexes and leading to the following sequence in periods 1 and 2:

1. Agents adjust downward their prior of a crisis at time 2 to $P_{1,2}$, $P_{0,1} \geq P_{1,2}$, anticipating that the actual probability of the crisis taking place in period 2 is $P_{1,2} \cdot Q(\rho_1)$. The updating follows the Bayesian inference outlined in the appendix.

2. The policy maker sets the intensity of regulation in period 1 at $\rho_1$. The income of the agents in periods 2 and 3 can be seen in table 1, updating the time indexes forward by one period.

3. At the beginning of period 2, the uncertainty regarding the crisis and its individual incidence is resolved. If no crisis takes place in period 2, the priors about the probability of a crisis in period 3 are updated as described above, and the problem repeats itself.

If a crisis takes place, a regime change may affect future regulations. In the absence of a regime change, the sequence described above continues with the proper adjustment of the prior regarding future crises. Agents are risk neutral, maximizing their expected utility, $\Sigma_{t=0}^{\infty} \beta^t C_t$. The recursive nature of the problem and the observation that regulation today affects the probability of a crisis tomorrow reduces the policy maker’s problem to a two-period optimization.
This section focuses first on the optimal regulation in period 0, contrasting two cases. The first is when the regulation intensity is determined by a planner maximizing the expected income of the “average agent”; the second is when the regulation is determined by a majority rule.

**Regulation Intensity Optimizing the Average Agent’s Welfare**

The policy maker sets the regulation intensity, \( r_0 \), to maximize the representative agent’s expected utility:

\[
\max r_0 \left[ 1 - \rho_0 \left[ (1 - \tau \rho_0) + \beta (1 - \tau P_{0,1} Q(\rho_0)) + (\omega + (1 - \omega) q) \right] \right].
\]

Equivalently, the regulation is maximizing:

\[
1 - \rho_0 + \beta [1 - \tau P_{0,1} Q(\rho_0)](\omega + (1 - \omega) q)\]

An internal solution with positive regulation will take place only if the marginal benefit of the first dollar spent on regulation exceeds 1: that is, if

\[
-1 - \tau \beta P_{0,1} Q(\rho_0) \cdot (\omega + (1 - \omega) q) > 0.
\]

Applying equation 1, \( Q(\rho_0) = -0.5 u \), positive regulation will take place only if \( 0.5 u \tau \beta P_{0,1} (\omega + (1 - \omega) q) > 1 \). High enough regulatory efficacy \( u \) will induce positive regulation. The critical level of regulatory efficacy associated with positive regulation will depend negatively on the probability of crisis \( P_{0,1} \); the cost of crisis, \( \tau \); the discount factor, \( \beta \); and the share of the agents exposed to the crisis, \( \omega + (1 - \omega) q \). Assuming an internal equilibrium, the first-order condition determining the level of regulation set by the planner, \( \tilde{\rho}_{s,0} \), is as follows:

\[
\tilde{\rho}_{s,0} = \begin{cases} \text{solution of} & 1 - P_{0,1} Q(\rho_0) \cdot \beta \tau \lambda_j - 1 \\ \text{if} & 0 < 0.5 u \beta \tau \lambda_j - 1 \\ 0 & \text{if} 0 \geq 0.5 u \beta \tau \lambda_j - 1, \end{cases}
\]

where \( \lambda_j = \omega + (1 - \omega) q \).

Optimal regulation equates a marginal dollar spent on regulation to its marginal benefit: the drop in the probability of a crisis times the discounted cost of a crisis, \( \beta \tau \), times the share of agents exposed to the crisis, \( \lambda_j: \lambda_j = \omega + (1 - \omega) q \).
**Regulation Intensity Set by Majority Rule**

The economy is composed of two groups; hence two cases correspond to this scenario. 

*The fully exposed (FE) group forms the majority,* as would be the case when $\omega > 0.5$. The regulation intensity is determined by applying considerations akin to the case of a planner, (equation 3), with the modification that the share of the agents exposed to the crisis increases from $\lambda_s = (1 - \omega) + \omega q$ to $\lambda_{FE} = 1$.

*The partially exposed (PE) group forms the majority,* as would be the case when individual uncertainty dominates, $\omega < 0.5$. The regulation intensity is determined in a way akin to the planner’s solution (equation 3), with the modification that the share of agents exposed to the crisis drops from $\lambda_s = (1 - \omega) + \omega q$ to $\lambda_{PE} = q$. Consequently, if an internal equilibrium takes place (that is, $\tilde{\rho}_0 > 0$), the optimal regulation equates the marginal cost, one, with the marginal benefit:

$$-P_{0,1} Q_{\rho}^\prime \beta \tau \lambda_k; \text{ where } \lambda_k = \begin{cases} \omega + (1 - \omega)q & \text{for } k = S \\ 1 & \text{for } k = FE \\ q & \text{for } k = PE \end{cases}$$

Consequently, the levels of regulation are ranked by claim 1.

**Claim 1**

With positive regulation, the level of regulation set by the “social planner” is below the level that is optimal for the FE group and above the level that is optimal for the PE group:

$$\tilde{\rho}_{FE,0} > \tilde{\rho}_{S,0} > \tilde{\rho}_{PE,0}.$$ 

The only difference between the first-order conditions determining optimal regulation is the probability of the deciding group, $\lambda$, being exposed to crisis incidence. If the regulation is set by the FE group, the probability is 1, as all agents in that group are fully exposed. If the regulation is determined by the PE group, this probability drops to $q$. The social planner attaches a probability of crisis incidence that is a weighted average of the incidence affecting the two groups: $\omega \cdot 1 + (1 - \omega)q$. As $q < \omega \cdot 1 + (1 - \omega)q < 1$, the weight attached to the marginal benefit of regulation is highest when the regulation is determined by a majority composed of the FE group, the lowest when the regulation is determined by a majority composed of the PE group, and in between these two when the regulation is determined by the planner.

Applying this discussion, the factors affecting the intensity of regulation are summarized in claim 2.

**Claim 2**

In an internal equilibrium with positive regulation, the greater the increase in the regulation rate the higher is the perceived probability of a crisis, the higher are the costs of a crisis, and the higher is the effectiveness of regulation, $\nu$. All of these factors also increase the likelihood of positive regulation.

To get further insight, this section reviews a simulation of an economy where individual-specific uncertainty dominates and the partially exposed agents are a majority, $\omega < 0.5$. Suppose that agents in the PE group are exposed to the crisis with
probability half \((q = 0.5)\) and that the conditions for internal equilibrium are met: 
\[ -1 + 0.5\beta v \tau q > 0. \]
Figure 4, panel a, illustrates a simulation where the share of the partially exposed agents is 0.7, and the probability \(P\) of a crisis in the absence of regulation, is half. The top dashed curve corresponds to the net marginal benefit of a social planner (\(S\)), and the solid curve corresponds to that of the \(PE\) decision maker. The optimal level of regulation is determined at the intersection of each curve with the horizontal axis. The \(PE\) group overlooks the regulations’ benefits accruing to the \(FE\) group. Consequently, the regulation set by the \(PE\) group is well below that of the social planner (\(\tilde{\rho}_{PE} = 0.026 < \tilde{\rho}_x = 0.0415\)). Figure 5 depicts the impact of the regulation on \(Q\): regulation intensity \(\tilde{\rho}_{PE} = 0.026\) cuts the probability of a crisis by about half, from \(0.5P\) to about \(0.25P\).
Intertemporal Inference and the Dynamics of Underregulation

This section evaluates the intertemporal patterns of underregulation during a spell of good times, a run with no financial crises. The appendix outlines a Bayesian inference example where the prior information about crisis occurrence follows a beta probability density function (pdf), with the prior mean of the crisis occurring in period 1 being $P_{0,1} = a / (a + b)$; $a, b > 0$. The coefficients $a, b$ reflect the prior information of the public regarding the mean and the variance of the probability of a crisis. As shown there, a run of good luck with no crises occurring in periods $(1, 2, \ldots, t)$ would induce a posterior mean of $P_{t, t+1} = a / (a + b + t)$. Hence a longer spell of good luck reduces the perceived mean of a crisis in the next period.

Consequently, each period that the crisis is avoided, the probability $P$ is adjusted downward, shifting both curves in figure 4 downward and thereby reducing the regulation intensity. A long enough spell of no crisis may induce zero regulation, as would be the case if $P$ drops from 0.5 to 0.25 in the above simulation. This situation is captured in panel b of figure 4. The socially optimal regulation is positive, yet the level of regulation set by the $PE$ group is zero: $\hat{\beta}_{PE} = 0 < \hat{\beta}_S = 0.02$. This simulation illustrates claim 3.

Claim 3

Suppose that the majority of the economy is partially exposed to the crisis. If the efficacy level of regulation is high enough, the level of regulation in period 0 is positive, but below the socially optimal level. The underregulation drops with the share of exposed agents, $q$. Tranquil time, when the crisis would not take place, reduces the intensity of regulation. If the spell of no crisis is long enough, the level of regulation may drop to zero, despite the fact that the socially optimal level remains positive.

Individual uncertainty regarding crisis incidences leads to underregulation, with the possibility of converging to zero regulation during long spells of no crisis, despite the fact that it is socially optimal to regulate. Consequently, uncertainty regarding the incidence of the crisis leads to underregulation, which increases the probability of the crisis above the socially optimum level. If regulation reduces both the probability of a crisis and the intensity of a realized crisis, then underregulation may increase the depth of the ultimate crisis. Claim 3 illustrates the paradox of prudential regulation: uncertainty regarding the identity of the agents that benefit from crisis avoidance leads to underregulation.

The impact of the precision of the prior information is summarized in claim 4, derived in the appendix.

Claim 4

The less informative is the prior regarding the probability of a crisis, the faster is the drop in regulations induced by a run of good luck.

With a less informative prior, the impact of a no-crisis event is greater, as the agent operates with greater initial ignorance. Thus good luck runs are especially
damaging in the context of financial innovations. Arguably, this was the case with some of the recent financial innovations in the United States (exotic derivatives, bundling). Therefore, good luck runs are more damaging in the aftermath of unprecedented financial reform, where the public is exposed to new financial instruments for the first time.

This discussion assumes that the regulatory effort is transparent and that the public is fully informed about it. The implications of regulation opacity and asymmetric information are summarized in claim 5 (see the appendix for the analysis).

**Claim 5**

Consider an economy with asymmetric information between the public and the regulator. The regulator determines the regulation intensity, but the public gets only noisy signals about it. A higher regulatory effort that avoids a crisis may be misconstrued as signaling that the environment is less risky, reducing the posterior probability of the crisis, $P_{t, t+1}$, below the level observed with symmetric information (recall that $P$ is the probability of the crisis in the absence of regulation). This may erode the future support for costly regulation.

This example deals with the costs of asymmetric information regarding the regulatory effort. Higher effort that helps to avert a crisis today would induce overconfidence, leading the public to infer that the risk is lower than the actual one, as the public underweighs the regulator’s effort. This in turn may reduce the support for costly future regulations and increase the ultimate cost of the crisis. Asymmetric information puts the regulator in the odd position of “damned if you do, damned if you don’t.” Probabilistically, with asymmetric information, the regulator would be damned if he is not putting in the effort today, as this may lead to a crisis tomorrow. Yet a regulator’s effort today may induce the public to be overly confident, as it undervalues the regulator’s contribution to crisis avoidance and overvalues the no-crisis event as a signal reducing the posterior probability of a crisis, $P$. Such overconfidence may lead to a deeper crisis in the future (see the appendix for a detailed example of this situation). Centralized regulation designed to reduce the confusion may mitigate the cost of asymmetric information.

**The Hazard of Overregulation**

The onset of a crisis may change this scenario in circumstances where the depth of the crisis exceeds the anticipated one. This may happen if the crisis would lead to the unexpected disappearance of markets for risk, implying that the realized cost of the crisis, $\tau$, is of a higher order of magnitude than the anticipated one. A possible interpretation of the disappearance of markets for risk is the emergence of Knightian uncertainty. This concept, postulated by Frank Knight, deals with situations in which agents who were exposed to quantifiable risks, drawn from known probability distribution, find that they operate in an environment where the probability distribution of a random outcome is unknown (what were perceived as “known unknowns” are
viewed now as “unknown unknowns”). The aversion to Knightian uncertainty may account for the flight to quality and the large cost of the current crisis, as well as for the costs of the 1997–98 East Asian crisis, when borrowers who were viewed as prime before the crisis lost access to credit.7

A crisis that leads to costs of a higher order of magnitude than the anticipated ones may induce a pendulum shift from underregulation to overregulation. Such overshooting was avoided in the 1990s. Indeed, the opposite took place in various countries (including Mexico and the Republic of Korea): following the onset of the crisis, countries adjusted by increasing their financial openness as part of a global trend of financial liberalization. This was probably because the crises in the 1990s were localized, at times of global growth, when the affected countries found that depreciation and higher exports facilitated the recovery. Today’s crisis is globalized, being propagated globally from the United States, leading to massive deleveraging of OECD exposure to developing countries. The global recession and the resulting drop in international trade imply that, on average, exporting the adjustment would not work this time. The present crisis exposes the fault lines of globalization, calling for domestic adjustment of policies and regulations in the OECD. While it is premature to reach a conclusion regarding the ultimate regulatory adjustment, the risk is that in the rush to regain credibility, policy makers may overshoot the adjustment or adjust in the wrong direction.

The argument about underregulation following a financial reform may have a symmetric counterpart dealing with overregulation following a financial calamity. In a system that represses financial intermediation, the stakeholders who would have benefited from financial intermediation are underrepresented in the decision making. This would be the case when there is individual uncertainty regarding the incidence of being a successful entrepreneur. The paper does not model this situation, as it is analogous to Fernandez and Rodrik (1991). The main difference between the case of underregulation and overregulation is that in the first, the absence of crisis induces a drop in the probability of a crisis, deepening the underregulation over time. In contrast, overregulation cuts the channels leading to the crisis, at a cost of reducing the actual output below the potential. Overregulation induces a static economy, where the benefit of avoiding a crisis may come with a larger cost of stagnation, a cost that may be underrepresented in the political discourse.

**On the Design of Balanced Regulations**

The paper closes with a discussion of regulatory changes needed to deal with the challenges associated with underregulation and overregulation.8

**Information Disclosure**

A necessary condition for successful regulation is that the regulator should be informed about the exposure to systemic risks. This requires having timely detailed information, preferably on a confidential basis, about financial institutions that are large enough to
be considered “too big to fail.” The lack of such information has been vividly illustrated in the United States.9 In the United States, the regulator (the federal government) imposes stringent disclosure requirements on the nonfinancial corporate sector, subject to strict confidentiality of the micro-level data disclosed to the regulator. Curiously, no comparable disclosure requirements are imposed on the financial sector. To illustrate, the Bureau of Economic Analysis (BEA) conducts an annual survey of U.S. direct investment abroad. The data collection is confidential and is based on mandatory surveys conducted by the BEA of all the establishments above a critical size. The survey contains detailed confidential information, including direct investment, employment data, research and development expenditures, trade in goods and services, and select financial data. This information, and other data collected by federal agencies, provides the regulator with timely information about the nonfinancial sector.

In contrast, in the decades before the crisis, there was no comparable attempt to collect data dealing with exposure of the financial sector. As a result, the regulator was frequently in the dark regarding the overall balance sheet exposure of investment banks, hedge funds, and other noncommercial financial intermediaries. This information gap is troublesome, as the cost of bailing out the financial system is frequently much higher than the cost of bailing out the nonfinancial real sector. Such cost discrepancy reflects the quick diffusion of financial panic. Furthermore, the financial sector tends to be more intertwined than the nonfinancial corporate sector, implying that the bankruptcy of a large enough financial institution may lead to large domino effects and systemic risks (see the massive bailout of AIG during the fall of 2008). Hence any serious regulatory reform should start with upgrading data collection, inducing mandatory periodic confidential reports of the balance sheet exposure of all financial institutions above a minimum size operating in the domestic market.

**Independence of the Regulatory Agency from the Political Process and Various Pressure Groups**

In the presence of individual risk regarding the incidence of a crisis, there are costs associated with designing regulations by a simple majority rule. Preferably, regulation should be the responsibility of an independent body managed by civil servants aiming at the expected GDP, with no sartorial biases. The logic for this independence is akin to the gains attributed to central bank independence. In the presence of principal-agent problems, the regulator’s independence is needed to avoid the wish of the regulated agents to minimize information disclosure. This follows from the moral hazard involved, where the agent would prefer underregulation as a way to facilitate excessive risk taking subsidized by the taxpayers. Indeed, the bargaining clout of Wall Street is seen as having contributed to the underregulation of the financial system in the United States.10

**Centralizing the Regulatory Process**

A fractured regulatory process runs the risk that each agency will focus on its narrowest task, viewing the need to deal with the big picture as beyond its mandate.
Facing the challenges of dealing with potential toxic assets, each regulatory agency would opt for a “not in my neighborhood” approach, preferring that other agencies deal with it. A fragmented regulatory approach is damaging, because with intertwining financial exposures, evaluating the systemic risk requires combining all the pieces of the financial puzzle together. A centralized regulatory process would also minimize the risk that a proactive financial system would attempt to arbitrage between diverse regulatory agencies.

Effective regulation also benefits by improving the quality of the signal provided by the regulatory agency. Fractured regulation runs the risk of increasing the opacity of the system, making it harder for the public to assess the regulatory effort. As shown in claim 5, greater opacity that leads to asymmetric information distorts the regulatory process and may magnify future underregulation.

Adopting Global Standards of Minimum Prudential Regulations and Information Disclosure, Enforced by the Domestic Regulator

The regulatory agency might face growing pressure to underregulate during spells of no crisis. Imposing global minimum standards increases the costs of deregulation, as it involves deviating from an international treaty. Thereby, such a treaty serves as a commitment device, increasing the odds of resisting transitory domestic pressure stemming from a run of good luck. Another rationale for the gains associated with minimum prudential standards follows the theory of the second best: the incidence of the moral hazard distortion associated with the “too big to fail” doctrine increases with the magnitude of the distorted activity. Underregulation, in a country that is financially open, may induce regulatory arbitrage, attracting capital inflows in search of higher returns induced by the higher implicit subsidy provided in more underregulated countries. Imposing minimum regulatory standards would help to mitigate the costs of such speculative inflows. Arguably, the large exposure of AIG to European and U.S. banks reflected such regulatory arbitrage: underregulation allowed AIG to sell underpriced insurance contracts that were de facto subsidized by the U.S. taxpayers. These contracts were in the form of a type of guarantee against default. As of June 2008, $307 billion of these contracts were written on instruments owned by banks in America and Europe and designed to guarantee the banks’ asset quality, thereby helping to boost their levels of regulatory capital. The sheer size of AIG made these contracts more attractive, increasing the odds of a bailout. Foreign parties operating in more conservative markets that did not offer such insurance (or markets that charged more for it) had the incentive to purchase the insurance offered by AIG, increasing the company’s short-run profits and the long-run costs to U.S. taxpayers.

Concluding Remarks

The framework outlined in this paper can be extended in various ways. The political equilibrium focused on the median voter in a model with two groups, assuming that
the median voter belongs to the partially exposed group. Similar results apply in a framework that allows for a large number of groups differing in their income and exposure to the crisis. As long as the median voter’s expected percentage income loss associated with the financial crisis is below that of the economy’s wide average, the outcome of the political equilibrium would be underregulation. Alternative modeling would allow for a lobbying equilibrium, where the policy is determined by the balance of contributions. Such modifications tend to magnify the underregulation in good times, as the stakeholders benefiting from underregulation may have access to deeper pools of funds supported by the rents associated with expected underregulation. Similar patterns would take place if agents’ priors are updated in a non-Bayesian manner, as may be the case if people pay insufficient attention to low-probability risks of disaster before a crisis happens and too much attention right after.13

The analysis of the paper suggests several testable implications. The conjecture that “longer runs” of good times are associated with faster deregulation may be tested directly, although one should recognize that deregulation may increase the duration of the run of good times, deepening the ultimate crisis. When the prior regarding the probability of crisis is less informative, the drop in regulation intensity associated with a spell of no-crisis events is faster. With less initial information, agents operating with greater initial ignorance put greater weight on new information. Thus good luck runs are more damaging in the aftermath of unprecedented financial reform, where the public is exposed to new financial instruments for the first time. This prediction may be tested, subject to the availability of indexes measuring the magnitude and the innovativeness of financial reforms and new financial instruments. Another finding of the paper is that with imperfect observability of the regulator’s acts, a higher regulatory effort would be confused with reduced risk in the system, reducing the agent’s posterior probability of a future crisis. This, in turn, leads to public complacency about regulation and to underinvestment in regulatory services. Measures of decentralization and opacity of regulatory efforts may be used to verify the degree to which this prediction accounts for the cross-country variation in exposure to the recent crisis as well as for the frequency and depth of crises over time.

In interpreting the results, one should keep in mind that the model applied in the paper is only one example of possible economic environments, where the duration of spells of good times affects the pattern of regulations. This model may be too limiting in that it does not consider the characteristics of the regulated entities or how they influence optimal regulation. Countries differ in structure, such as managerial capacities, political instability, polarization, and aversion to downside risk. These differences may be manifested in different regulatory structures. Yet, independently of the structure, the regulators may face similar challenges across countries, adjusting regulation to reduce exposure to systemic risk, while not forcing banks to become too conservative, and to produce a level of risk that is below the socially optimal level. It is important to keep in mind the difference between regular regulation and systemic regulation. The latter has received much attention during this crisis period. It appears that the costs and benefits associated with crisis occurrence in an uncertain environment is only one dimension of the bigger picture dealing with the short- and long-run trade-offs between a “regular” and a “systemic” regulatory structure.
To conclude, the challenge facing the global financial system is to reform the global financial architecture to allow Goldilocks regulations, mitigating the temptation to underregulate in spells of good times and preventing overregulation in the aftermath of a financial crisis. The risk of not meeting these challenges is that some affected countries will opt to reduce their financial integration, some will overshoot the regulatory adjustment and thereby compromise future growth, and some will remain exposed to the hazard of replaying crisis dynamics in the future.

Appendix

This appendix reviews a Bayesian inference problem leading to the results discussed in the paper (see Zellner 1971 for detailed analysis of similar inference problems). Assume that the probability of the crisis is constant over time. Hence the probability of \( n \) independent draws from the crisis distribution follows a discrete binomial pdf. Specifically, the probability of observing during \( n \) periods \((1, 2, \ldots, n)\) of which \( m \) tranquil, no-crisis events and \( n-m \) crisis events is

\[
p(m \mid \theta, n) = \binom{n}{m} \theta^m (1-\theta)^{n-m}
\]

Equation A.1, viewed as a function of the unknown parameter \( \theta \), is the likelihood function. Suppose that the prior information about the probability of crisis occurrence, \( \theta \), is summarized as in the following beta pdf:

\[
p(\theta) = k \theta^{a-1} (1-\theta)^{b-1}; \quad a, b > 0
\]

where \( k \) is a normalization constant, \( k = \Gamma(a + b) / \Gamma(a) \Gamma(b) \). Applying the properties of the beta pdf, it follows that the prior mean of the crisis in period 1 (\( P_{0,1} \) in the notation of the paper, for the case where \( Q = 1 \)) is

\[
E(\theta) = \frac{a}{a + b}
\]

It can be shown that \( V(\theta) = ab / [(a + b)^2 (a + b + 1)] \). Hence the values of \( a \) and \( b \) reflect the public’s prior information about the mean and the variance of crises occurrence.

Applying equations A.1 and A.2 and Bayes’s theorem, the posterior pdf for the crisis after \( n \) periods with \( m \) crisis realizations is a modified beta pdf:

\[
p(\theta \mid n, m) = k' \theta^{a-1+m} (1-\theta)^{b-1+n-m}; \quad a, b > 0
\]

where \( k' \) is the modified normalization constant. The posterior mean of a crisis can be shown to be

\[
E(\theta \mid n, m) = \frac{a + m}{a + b + n}
\]

Thus no crisis in period 1 results in a posterior probability of a crisis in period 2 of \( P_{1,2} = a / (a + b + 1) \). Equation A.5 also implies that the posterior mean of a crisis after a good luck spell of no crisis during period 1, \( \ldots, t \) is \( E(\theta \mid t) = a / (a + b + t) \). Hence the marginal impact of another period of no crisis on the crisis probability is

\[
\frac{\partial P_{1,t+1}}{\partial t} = - \frac{P_{1,t+1}}{a + b + c}
\]
Recalling that $V(\theta) = ab / [(a + b)^2 (a + b + 1)]$, more informative priors are associated with higher $a$ and $b$ $[V(\theta) \to \infty, \theta \in [0, \infty])$. Thus equation A.6 implies that less informative priors would increase the impact of another period with no crisis on the posterior probability of a crisis (claim 4 in the main text).

This appendix closes with an overview of claim 5. Suppose that the public has incomplete information about regulatory efficacy. This would be the case if the actual $Q(\rho_0)$ is affected by regulatory effort, which is noisily observed by the public. For example, suppose that the actual impact of the regulation intensity increases with the regulatory effort, $e$:

$$\tilde{Q}(\rho_0, e) = Q(\rho_0)(1 - e),$$  \hspace{1cm} (A.7)

where $\tilde{Q}$ is the regulatory impact on $P$, as a function of regulatory cost and effort, $\rho_0$, $e$, respectively. The public observes the cost of regulation, $\rho_0$, but has only a noisy indicator $\varepsilon$ of the regulator’s effort, $\varepsilon = e + \delta$, where $\delta$ is the noise. The two components of $\varepsilon$ are assumed to be zero mean, independent, $E(\varepsilon) = E(\delta) = \text{cov}(\varepsilon, \delta) = 0$, and uncorrelated with the factors affecting the occurrence of a crisis in the absence of regulation.

In these circumstances, observing $\varepsilon$ leads the public to infer that the expected regulatory effort is

$$E(e | \varepsilon) = \varepsilon \frac{V(e)}{V(e) + V(\delta)}. \hspace{1cm} (A.8)$$

Higher noisiness of the information, that is, higher $V(\delta) / V(e)$, implies that a given observed $\varepsilon$ is associated with lower perceived regulatory effort. Observing $\varepsilon$, the public infers that the regulation function is

$$\tilde{Q}^p(\rho_0, e) = Q(\rho_0) \left(1 - e \frac{V(e)}{V(e) + V(\delta)}\right). \hspace{1cm} (A.9)$$

The actual “regulation function” observed by the regulator is $\tilde{Q}^a(\rho_0, e) = Q(\rho_0)(1 - e)$, where the upper indexes $p$ and $a$ indicate the public inferred $\tilde{Q}$ and actual $Q$, respectively. Thus the public underestimates the impact of the regulator’s effort by

$$\tilde{Q}^a(\rho_0, e) - \tilde{Q}^p(\rho_0, e) = Q(\rho_0) \left[\delta \frac{V(e)}{V(e) + V(\delta)} - e \frac{V(\delta)}{V(e) + V(\delta)}\right]. \hspace{1cm} (A.10)$$

The public’s undervaluation of the regulatory effort increases with the noisiness of the information, that is, $V(\delta) / V(e)$, and with the realized effort $e$. Hence greater regulatory effort that prevents a crisis would induce the public to understate the role of the regulator, magnifying the reduction in the perceived risk of a crisis. The asymmetric information about the actual effort implies that if no crisis takes place in period 1, the posterior probability of a crisis in period 2 will be lower than the one with symmetric information.

To illustrate, suppose that $e$ and $\delta$ follow the simplest discrete distribution, each having values $(h, -h)$, with probability half. Hence the observed $\varepsilon$ has three possible values: $2h; -2h$, corresponding to $(e, \delta) = (h, h)$ or $(-h, -h)$, each pair having a probability 0.25;
and 0, with probability half, corresponding to \((e, \delta) = (h, -h)\) or \((-h, h)\), each pair having a probability 0.25. Suppose that the regulator chooses the high effort \(e = h\). With probability half, \(\delta = -h\), hence \(e = 0\), inducing the public to infer correctly the regulatory effort. With probability half, \(\delta = h\), hence \(e = 2h\). Not having the full information about \(e\) and \(\delta\), the public notes that this signal is consistent with either high or low effort \((e = h; \delta = -h\) or \(e = -h; \delta = -h)\), each with probability half. Thus the public’s inferred expected effort, conditional on \(e = 0\), is \(E(e | e = 0) = 0\).

Consequently, had the regulator followed the high-effort enforcement \((e = h)\), the public would infer expected effort of \(0.5h\) \([= (0.5 \times 1 + 0.5 \times 0)h]\), instead of the actual effort, \(h\). The gap between the actual and the expected effort is the outcome of asymmetric information. Consequently, had the regulator followed the high-effort policy \((e = h)\), the public’s expected posterior in the absence of the crisis would be lower than the symmetric information, \(P_{1,2} = a/(a + b + 1)\), in states when the realized signal would be \(e = 0\).

This example illustrates the distortion induced by asymmetric information regarding the regulatory effort. Higher effort that helps to avert a crisis today would induce overconfidence, leading the public to infer that the risk is lower than the actual one, as the public underweighs the role of the regulator. This, in turn, would reduce the support for regulations in the future and might increase the ultimate cost of the crisis.

Notes

1. The choice of a de jure measure of capital account openness is driven by the motivation to look into the policy intentions of the countries. The Chinn and Ito index (Ito and Chinn 2008) is based on information in the International Monetary Fund’s Annual Report on Exchange Arrangements and Exchange Restrictions (IMF various years). Specifically, it is the first standardized principal component of the variables that indicate the presence of multiple exchange rates, restrictions on current account transactions and on capital account transactions, and the requirement of the surrender of export proceeds. Higher values of this index indicate that a country is more open to cross-border capital transactions. The index is available for 171 countries for the period of 1970 through 2006.

2. See Aizenman and Noy (2009) for further analysis of such two-way positive feedback effects, applying panel regressions and Geweke’s decomposition methodology.


4. “As we move into a new century, the market-stabilizing private regulatory forces should gradually displace many cumbersome, increasingly ineffective government structures.” Remarks by Alan Greenspan, April 12, 1997.

5. The marginal gain associated with increasing the regulation from the perspective of the planner is the bold, top curve, plotting \(\beta(-\tau_{0,1} \cdot P_{0,1} \cdot Q' \rho_0 [\omega + (1 - \omega)q] - 1\). The mar-
ginal gain associated with increasing the regulation from the perspective of the PE group is the lower curve, plotting $\beta(-\tau P_{0,1} \cdot Q_{P_0}(q) - 1$. The social marginal benefit of regulation exceeds the marginal benefit assessed by the PE group, as $\omega + (1 - \omega)q > q$.

6. The vertical gap between the two curves, $-\tau P_{0,1} \cdot Q_{P_0}(1 - q)$, drops with the share of exposed agents in the PE group. Thus individual uncertainty is a key factor accounting for the gap between the socially optimal level of regulation and the regulation determined by group PE; if all agents are fully exposed ($q = 1$), the majority would choose the socially optimal level of regulation.

7. See Caballero and Krishnamurthy (2008) and Blanchard’s guest article in the Economist (Blanchard 2009) for the role of Knightian uncertainty in explaining flight to quality and the disappearance of risk markets.

8. See Barth and others (2009) for a comprehensive empirical study of the impact of differential bank regulations and supervision.

9. The crisis exposed the large systemic exposure of the U.S. financial system to credit-default swaps at levels that surprised both the public and the regulators (see the case of AIG).

10. See Rajan and Zingales (2003) and Rajan (2005) for insightful discussions of the political economy aspects of financial intermediation. They point out the hazard associated with the incumbent’s ability to leverage the power of government regulation to protect the incumbent’s economic position. This comes frequently at the expense of the public interest. Similar concerns were raised recently in the United States: “The New York Fed is, by custom and design, clubby and opaque. It is charged with curbing banks’ risky impulses, yet its president is selected by and reports to a board dominated by the chief executives of some of those same banks.” Jo Becker and Gretchen Morgenson, “Geithner, Member and Overseer of Financial Club,” New York Times, April 26, 2009.

11. For further discussion, see Aizenman (2009).


References


Barth, R. James, Chen Lin, Yue Ma, Jesus Seade, and Frank M. Song. 2009. “The Role of Bank Regulation, Supervision, and Monitoring in Bank Efficiency.” Unpublished mss, Auburn University, Auburn, AL.


Aizenman’s paper addresses a timely and important topic. The global financial crisis has prompted a reassessment of financial regulatory systems worldwide. Financial crises by their very nature accentuate the need for (usually overdue) regulatory reform. True to form, the current turmoil exposed shortcomings in supervisory, regulatory, and prudential frameworks (see ADB 2009), leading national authorities—together with regional and global financial institutions—to reexamine approaches to financial regulation and supervisory oversight. How should the global regulatory framework be reconstructed to safeguard the stability of financial systems, prevent the next crisis, and ensure sustained, robust global economic growth?

Aizenman’s paper focuses on the procyclicality and suboptimality of the current regulatory structure. Extended periods of economic tranquility and financial stability tend to reduce the demand for regulation, leading to underregulation. Similarly, there is the risk of overregulation following financial crises. Hence regulation tends to be procyclical.

Aizenman proposes a regulatory structure that can mitigate the suboptimality or procyclicality. To ensure a sufficient degree of regulation “through the cycle,” the paper argues, regulatory reforms should improve information disclosure, increase the regulator’s independence, centralize the regulatory process, and build global standards of prudential regulation and information disclosure.

The paper presents a model to show an imbalance between individual demand for regulation and the socially optimal level of regulation supported by the majority. The model emphasizes that financial reforms in the presence of individual-specific uncertainty could lead to a level of regulation below what is socially optimal in good times.
leading up to a crisis. In the wake of a crisis, the crisis response then brings on over-regulation. The simple political economy model draws this important prediction based on several assumptions: (a) regulation is costly, (b) agents have individual risks and update their assessments of the probability of crisis, and (c) regulation intensity (spending resources) is set by majority rule.

The model is innovative, offering interesting and paradoxical results that show financial reforms tend to add procyclicality to regulation. Some aspects of the model, however, remain unclear. Is it really true that underregulation is mainly an outcome of a political process where the majority rules? Can an omnipotent central regulator be a social planner who knows the socially optimal level of regulatory intensity? And, in practice, who can be that social planner or omnipotent central regulator?

In my view, the model does not fully account for the inadequate institutional capacity of the current financial regulatory structure to keep in step with rapid financial innovation and globalization. This is important because it raises two issues: (a) how to address systemic failures in the financial system and (b) how to build a new financial regulatory framework.

**Systemic Failures of the Financial System**

The global financial crisis exposed two major weaknesses in the current regulatory and supervisory structure: first, the regulatory structure was unable to deal with the innate procyclicality of financial systems, and, second, the lack of systemwide macroprudential oversight allowed systemic failures to occur in the financial system.

First, the current crisis highlighted several mechanisms that contributed to procyclicality in the financial system. Many factors influencing private sector behavior and practices, prudential regulation, and macroeconomic policies accentuate cyclical movements in the financial system.

The mechanisms that contribute to procyclicality in market and credit risk management systems include (a) use of the value-at-risk model, which encourages firms to increase their risk appetite in low-volatility environments and to reduce it in high-volatility environments, (b) credit ratings that are also procyclical, and (c) compensation practices that reward managers for taking excessive risks for short-term returns rather than for taking a longer-term view of business prospects and their associated risks.

The existing regulatory system inadequately addresses procyclicality and, in some cases, even encourages procyclical tendencies. The crisis revealed incidences where regulation, supervision, and risk management failed to tame excess leverage and poor risk management or to correct the flawed incentive structures of financial institutions. Procyclical regulation of bank capital, short time horizons in risk assessment, and fair value accounting contributed to the procyclicality of regulatory frameworks (see Andritzky and others 2009).

Second, the aim of macroprudential supervision is to ensure the stability of the financial system by appropriately monitoring all financial activities that may pose
systematic risks. The existing regulatory framework, which emphasizes microprudential supervision, was unable to identify the buildup of systematic risks. In hindsight, aggregate shocks were underestimated.

The call for stronger macroprudential capabilities in no way implies that microprudential measures are wrong or no longer needed. Nevertheless, traditional regulation, which focuses almost exclusively on individual institutions and specific financial instruments, will likely fall short in providing effective regulation for increasingly interdependent financial institutions and markets. Effective regulation should take appropriate account of systemic and cyclical factors to create awareness of overall systemic leverage and to mitigate the potential procyclical effects of regulation.

How to Build a New Financial Regulatory Framework

We all agree that regulation of financial markets and institutions must be overhauled. The principal goals are to promote more robust risk management and to establish more effective prudential oversight.

First, we need a systematic approach to build up an adequate framework of macroprudential oversight that can counter the procyclical effects of prudential regulations. New mechanisms include (a) countercyclical capital regulations, (b) countercyclical loan-loss provisioning requirements, (c) measures limiting the procyclicality of property lending and unhedged foreign currency credit, (d) measures limiting foreign exchange risks by imposing limits on foreign exchange exposure, (e) more intensive monitoring of problem financial institutions, and (f) better information disclosure.

Second, there should be a balance between innovation and regulation. The crisis taught us that market discipline may not be a perfect substitute for prudential regulation. But we must also be mindful of the risks of overregulation when revamping the current regulatory system. Regulation should not constrain financial innovation. A strong institutional framework must ensure that the functioning of a financial system is aligned with appropriate market incentives—a key element for sustainable regulation. Market discipline also needs to be strengthened by improving transparency and creating more incentive-compatible compensation structures.

Third, there is no single international best practice, especially in constructing an effective regulatory environment. Regulatory costs and benefits change across economies and over time, depending on the structural and institutional characteristics of national financial systems. A regulatory structure should provide competitive neutrality, avoid duplication of scarce supervisory resources, and effectively address the issue of institutional solvency. Still, there is no one-size-fits-all regulation.

Fourth, national enforcement and global coordination should come together. Policy responses should be coordinated internationally to avoid regulatory arbitrage and competitive distortions. Financial integration limits the effectiveness of unilateral measures. Measures will be more effective if supervisory agencies collaborate closely (see ADB 2009).
There are practical impediments to institutionalizing a global regulatory agency. Therefore, existing global financial institutions—such as the Financial Stability Board or the International Monetary Fund—have important roles to play. We should promote global coordination in sharing information, employing international accounting standards, and harmonizing minimum prudential rules.

It is clearly in Asia’s interest to be an active participant in efforts to design a global financial architecture that meets the challenges of globalized finance. We must work together in constructing a new global financial architecture.

References


The Road Ahead to a Sustainable Global Economic System
Lessons from the Recent Financial Crisis for Reforming National and International Financial Systems: The Road Ahead to a Sustainable Global Economic System

STIJN CLAESSENS

The recent financial crisis is due to a confluence of many well-known factors and some new ones. It has brought to light many weak elements in national financial architectures, particularly in the treatment of systemic banks and other financial institutions, the assessment of risks and vulnerabilities, and the resolution of financial institutions and claims. The global nature of the financial crisis has shown that financially integrated markets have benefits, but also risks, with real economic consequences. It has shown that the international financial architecture is still far from matching the world’s closely integrated financial systems. The crisis has had major financial and economic repercussions for emerging markets and developing countries alike. Countries are benefiting from their improved fundamentals. To tackle the crisis, short-term policy responses, involving more accommodative fiscal and monetary policy and better restructuring frameworks, are being put in place. But the crisis also highlights specific challenges for financial sector reform for emerging markets and developing countries.

This paper draws lessons from the recent financial crisis for reforming national and international financial systems, including short-term policy implications for emerging markets and developing countries. To diagnose the problem, the paper starts with a review of the causes of the current global financial crisis that draws on historical perspectives and discusses its international dimensions. It highlights the multiple causes of the crisis, with a mixture of some elements common to other financial crises and some new elements. It reviews the many channels and mechanisms through which the...
financial crisis propagated and spread globally. And it shows how the ongoing global crisis is leaving a considerable legacy of government interventions and macroeconomic consequences, especially in advanced countries, that will condition future actions and reforms.

The paper then reviews the lessons for national and international financial reforms. The financial crisis has brought to light many weak elements in the financial architecture of nations, particularly regarding the treatment of systemic banks and other financial institutions, the assessment of risks and vulnerabilities, and the framework for resolving financial institutions and claims. The global nature of the financial crisis has shown that financially integrated markets have benefits, but also risks, with large real economic consequences. The crisis highlights that the international financial architecture is still far from matching the world’s closely integrated financial systems. Improvement is needed in the areas of surveillance, information sharing, crisis management, and liquidity support. The paper summarizes current thinking on what reforms can best address these issues.

Particular emphasis is then given to the implications and lessons of this crisis for emerging markets and developing countries. In the short run, policy makers in these countries have to deal with the financial and economic consequences of a crisis often not of their making. But many countries are better positioned than in the past to deal with these challenges, given their improved fiscal positions, stronger financial sectors, and improved institutional environments. Still, these countries face numerous challenges highlighted by the financial crisis. The paper concludes with several areas of current debate and areas where more research would be useful to help guide policy makers.

Causes and Evolution of the Crisis and the Current State of Affairs

This section reviews the causes of the financial crisis, including a mixture of some elements common to other financial crises and some new elements. It then reviews the channels and mechanisms through which the financial crisis propagated and spread, showing that, while the crisis emerged in the U.S. housing markets, it quickly broadened to other financial markets in the United States and globally. Finally, it reviews how the crisis has left a considerable legacy of government interventions and macroeconomic consequences, which will condition future actions and reforms.

The 2007–08 Financial Crisis and Other Crises: Similar, Yet Different?

The severe financial crisis that has gripped the global economy reflects a remarkable confluence of factors. Some are reminiscent of past bouts of financial turmoil, but others are new (and surprising). This section identifies both what is common and what is different between the current crisis and previous ones. While ranking the relative contributions of the causes of the crisis is not without controversy, together these elements help to explain the considerable scale and scope of the current
Commonalities with Previous Crises

The crisis shares four features in common with other crises: asset price increases that were not sustainable, credit booms that led to excessive debt burdens, the buildup of marginal loans and systemic risk, and the failure of regulation and supervision to keep up with and get ahead of the crisis when it erupted.

Housing prices rose sharply in the United States and other markets prior to the current crisis (see figure 1). The patterns of asset prices are reminiscent of those in other major financial crises (see figure 2). The overall size of the U.S. housing boom and its dynamics—including a rise in housing prices in excess of 30 percent in the five years preceding the crisis and peaking six quarters prior to the beginning of the crisis—are remarkably similar to the development of housing prices in the previous (Big Five) banking crises in advanced economies (Finland, 1991; Japan, 1992; Norway, 1987; Sweden, 1991; and Spain 1977; see Reinhart and Rogoff 2008).

Such sharp increases in housing prices were also common in other countries hard-hit by the current crisis and were associated with rapid growth in credit aggregates (see figure 3). Housing prices rose rapidly in many countries now caught in the financial turmoil, including the United Kingdom and Iceland. These housing booms were generally fueled by fast-rising credit resulting in sharply higher household leverage.
The prolonged credit expansion in the run-up to the crisis is similar to other episodes (see figure 4). Sustained episodes of rapid credit growth generally coincide with large cyclical fluctuations in economic activity—with real output, consumption, and investment rising above trend during the buildup phase of credit booms and falling below trend in the unwinding phase (Mendoza and Terrones 2008). In the upswing, the current account tends to deteriorate, often accompanied by a surge
in private capital inflows. Increases in housing prices and the real exchange rate often accompany such credit booms.

While aggregate credit growth in the United States was less pronounced than in previous episodes, reflecting slower expansion of corporate credit, household debt increased sharply. Credit to households rose rapidly after 2000, driven largely by outstanding mortgages, interest rates below historic averages, and financial innovation. And in spite of low interest rates, debt service relative to disposable income reached a historic high. The increased leverage left households vulnerable to a decline in housing prices, a tightening in credit conditions, and a slowdown in economic activity. Similar patterns existed in several crisis countries.

As in other crises, the rapid expansion of credit seems to have played a role in the current crisis. While historically only some credit booms end up in financial crisis, the probability of a crisis increases with a boom (see figure 5; see Dell’Ariccia, Barajas, and Levchenko 2009). Furthermore, the larger are the size and duration of a boom, the greater is the likelihood that it will result in a crisis. The mechanisms linking credit booms to crises include an increase in leverage of borrowers (and lenders) and a decline in lending standards. In the U.S. episode, both channels were at work (see figure 6; see Dell’Ariccia, Igan, and Laeven 2008).

This pattern extended to various extents to other countries caught in the current storm (see figure 7). In the run-up to the crisis, credit aggregates grew very quickly in the United Kingdom, Spain, Iceland, and several Eastern European countries. As in the United States, these credit expansions often fueled real estate booms. Increased international financial integration supported these patterns. For many countries, a clear relationship existed between credit growth and capital inflows (see figure 8).
The boom in household credit was associated with the creation of marginal assets whose viability relied on continued favorable macroeconomic conditions. In the United States (and to some extent the United Kingdom), a large portion of the mortgage expansion consisted of loans extended to subprime borrowers with limited credit and employment histories (see figure 9). Debt servicing and repayment were, hence, vulnerable to economic downturns and changes in credit and monetary conditions. This maximized default correlations across loans, generating portfolios
highly exposed to declines in housing prices, confirmed ex post through the large nonperforming loans when house prices declined.

Elsewhere, a similar pattern led to large portions of domestic credit denominated in foreign currency. Large foreign currency exposures in the corporate and financial sectors were common in the Asian crisis. In the current crisis, in several Eastern European economies large portions of domestic credit (including to households) were denominated in foreign currency (euros, Swiss francs, and yen; Árvali, Driessen, and Ötker 2009). While lower interest rates on foreign currency loans relative to local currency loans increased affordability, borrowers' ability to
service loans and creditworthiness depended on a stable exchange rate. As with U.S. subprime loans, this meant highly correlated default risks across loans, which exposed the system to macroeconomic shocks.

On the back of buoyant housing and corporate financing markets, many forms of derivatives markets expanded greatly. Favorable conditions spurred the emergence of large-scale derivatives markets, such as mortgage-backed securities and collateralized debt obligations with payoffs that depended in complex ways on underlying asset prices. The pricing of these instruments was often based on a continuation of rising housing prices, which facilitated the refinancing of underlying mortgages. The corporate credit-default swap market also expanded dramatically on the back of favorable spreads and low volatility.

Past crises often followed expansions triggered by financial liberalization not accompanied by necessary regulatory reforms. Imbalances often were the result of badly sequenced regulatory reforms. Poorly developed domestic financial systems were often unable to intermediate large capital inflows in the wake of capital account liberalizations. Poorly designed financial reforms and deficient supervision often led to currency and maturity mismatches and to large, concentrated credit risks.

In this crisis, although perhaps in more subtle forms, regulatory approaches to and prudential oversight of financial innovation were insufficient as well. As in the past, but this time in advanced countries, finance companies, merchant banks, investment banks, and off-balance-sheet vehicles of commercial banks operated—to varying degrees—outside banking regulations. Providing increasingly important avenues for intermediation, this “shadow banking system” grew without adequate oversight, leading to systemic risks. Regulators also underestimated the conflict of interests and information problems associated with the originate-to-distribute model. Not only did this harm consumers of financial services, but it also created the potential for chain reactions leading to systemic risk.

As often before, the focus of authorities remained primarily on the liquidity and insolvency of individual institutions, rather than on the resilience of the whole financial system. This meant an underestimation of the probability and costs of systemic risk. At
the international level, insufficient coordination among regulators and supervisors and the absence of clear procedures for resolving global financial institutions hindered efforts to prevent and contain the impact and transmission of the crisis.

And, in terms of crisis response, as in past events, it has proven difficult to get ahead of a fast-evolving situation to contain the financial turmoil and reduce the impact on the real economy. Ad hoc and piecemeal interventions created further disruptions and loss of confidence among creditors and investors. The chronology of the crisis (Calomiris 2009; Gorton 2008) shows how events and market developments triggered and conditioned subsequent developments and policy responses that, in retrospect at least, made the crisis more severe.

**New Dimensions of the Crisis**

New dimensions played an important role in the severity and global scale of the crisis, particularly with respect to its transmission and amplification. Four key aspects were new: the widespread use of complex and opaque financial instruments; the increased interconnectedness among financial markets, nationally and internationally, with the United States at the core; the high degree of leverage of financial institutions; and the central role of the household sector.

Securitization and innovative (but complex) financial instruments were a critical element of the credit expansion in U.S. securitization. Although securitization has been a long-standing technique for prime loans conforming to the underwriting standards of government-sponsored enterprises, in 2007 more than 70 percent of nonconforming mortgages in the United States were securitized, up from less than 35 percent in 2000 (see Ashcraft and Schuermann 2008; Gorton 2008; Brunnermeier 2009). Other assets were increasingly packaged as well, and cash-flow streams from securities were further separated and tranched into other securities (for example, collateralized debt obligations).

In part by being inadequately regulated, the increased recourse to securitization and the expansion of the originate-and-distribute model exacerbated agency problems. The progressive expansion of more opaque and complex securities and the increasing delinking of borrowers from lenders worsened agency problems. Risk assignments became increasingly unclear, and incentives for due diligence worsened, leading to insufficient monitoring of loan originators and an emphasis on boosting volume to generate fees. The distribution model led to widespread reliance on ratings for the pricing of credit risks, with investors often unable or unwilling to assess underlying values and risks.

Increased balance sheet opaqueness and reliance on wholesale funding increased the fragility of the system. Once U.S. housing prices began to decline and defaults began to rise (affecting the expected value of the assets underlying market-backed securities and collateralized debt obligations), the complexity of instruments undermined price discovery and led to market illiquidity and a freeze on securitization activity. The increased opaqueness of balance sheets (due to the widespread recourse to off-balance-sheet instruments) made it difficult to separate healthy from unhealthy institutions. The resulting adverse selection problems contributed to the freezing of interbank markets and forced further sales of securities to raise funds.
The increased centrality and systemic importance in many countries of highly leveraged, underregulated intermediaries relying on wholesale and short-term funding exacerbated problems.

Financial integration rose dramatically over the past decade, especially among advanced economies. Capital account openness and financial market reforms led to massive increases in cross-border gross positions, especially among high-income countries (see figure 10). The presence of foreign intermediaries also increased in several banking systems (including in many emerging markets). As a result, while international risk sharing, competition, and efficiency increased, so did the risk of transmitting financial shocks across borders.

Increasing interconnectedness of financial institutions and markets (see figure 11) and more highly correlated financial risks intensified cross-border spillovers through many channels, including liquidity pressures, global sell-off in equities (particularly financial stocks), and depletion of bank capital. Mortgage-backed securities were widely held by institutions in the United States, but also in other advanced economies, and by the official sector in several emerging markets. As troubled intermediaries hit by losses and scrambling for liquidity were forced to sell other assets and cut lending, the crisis gradually spread to other markets and institutions through “common lender effects.” Emerging markets—especially those that had relied heavily on external financing and, paradoxically, those with more liquid markets—were affected through capital account and bank funding pressures.

The sheer size of the U.S. financial market and its central role as an investment destination contributed to the spread of the crisis. Any shock to the U.S. financial markets and economy is bound to have global effects. U.S. financial assets repre-
sent about 31 percent of global financial assets, and the U.S. dollar share in reserve currency assets is about 62 percent. In recent years especially, U.S. financial assets were perceived to offer the combination of safety and liquidity attractive to private and public investors alike. More generally, since the United States is a large economy, it has a large effect on global developments. The diversified structure of international financial markets made coordination difficult.

The crisis also triggered the unwinding of imbalances in other countries. Benign financial and macroeconomic conditions—notably, low interest rates and narrow risk spreads—occurred on a global basis, fueling booms in many economies. Housing market vulnerabilities came home to roost in several countries, notably in Europe. In the United Kingdom, with a similar housing boom as in the United States, mortgage lenders came under intense pressure, beginning in fall 2007. Large pressures hit Iceland, Hungary, and the Baltic countries where imbalances were pronounced. The increased connections and simultaneous buildup of systemic risks across multiple countries made the management of shocks more complex, especially

FIGURE 11.
Growth in International Claims and Foreign Exposure, Various Years

Source: Bank for International Settlements.

a. Excludes foreign currency claims on home-country residents. Nordic countries include Danish, Finnish, Norwegian, and Swedish banks. Other nationalities include total international claims, excluding those booked by Japanese, Nordic, and U.S. banks.

b. On an ultimate risk basis and excluding interoffice transfers. Foreign claims vis-à-vis entities (banks and non-banks) in advanced economies, booked by banks headquartered in the countries shown. Total is on the left-hand scale; individual countries are on the right-hand scale.
in light of institutional deficiencies in many countries, including the inability to resolve large, cross-border financial institutions quickly.

The buildup of an unusually high degree of leverage of financial institutions and borrowers contributed to the propagation of shocks. Leverage increased sharply in the financial sector, both directly at commercial banks in Europe and indirectly through the shadow banking system and the rising share of investment banks and non-deposit-taking institutions in the United States (see figure 12). This buildup of leverage among households was notably different from the situation in previous crises. In the run-up to Japan’s real estate crisis, for example, while the household debt-to-income ratio rose sharply, measures of household leverage (household debt-to-assets ratio) declined, suggesting that Japanese homeowners built equity in their properties as real estate prices soared.

This high leverage limited the system’s ability to absorb even small losses and contributed to the rapid decline in confidence and increase in counterparty risk early in the crisis. Loan-to-income values larger than in the past left households highly exposed to shocks, while at the same time high loan-to-value mortgages allowed even moderate declines in housing prices to push many households into negative equity. In the financial sector, initial concerns about liquidity could give way quickly to worries about solvency.

The buildup in leverage (including rising household indebtedness) was not restricted to advanced economies. In some emerging economies, vulnerabilities were related to rising reliance on external financing flows. Amid global deleveraging, heightened investor risk aversion, and repatriation of funds, many emerging economies suddenly found foreign funding sources increasingly scarce and were
confronted with sudden stops or reversals of capital flows. In addition, emerging-market corporations faced much higher borrowing costs, limited opportunities to issue equity, and few alternative sources of financing. While official financing filled some of the gaps, emerging markets had to make rapid adjustments, leading to real economic dislocations.

Problems in the household sector have played a more prominent role in this crisis than in previous crises. Most previous episodes of financial distress stemmed from problems in the official sector (for example, Latin America’s debt crisis of the 1980s) or the corporate sector (for example, the Asian crisis of the late 1990s). The current crisis, however, largely originated from overextended households, in particular from subprime mortgage loans (see figure 13). This had implications for how the crisis was transmitted from the financial to the real sector and complicates the resolution mechanisms and policy responses.

In the United States, a vicious cycle of rising foreclosures, falling home values, and disappearing securitization markets quickly developed. Vulnerable cohorts of
borrowers became increasingly susceptible to rising interest rates and falling home values and could no longer refinance their mortgages, leading to higher monthly payments, rising delinquencies, and rising default rates. A wave of finance company failures—suddenly no longer able to securitize subprime mortgages—led to a breakdown in mortgage origination and an abrupt adjustment. Adverse feedback loops—rising foreclosures placing additional downward pressures on housing prices—started. With U.S. housing prices declining on a national basis for the first time since the Great Depression, many heavily indebted borrowers confronted with substantial negative home equity faced incentives to “walk away.”

Tightening standards for new mortgages and consumer credit led to a sharp compression in consumer spending that compounded difficulties in the real sector. With households’ savings and net assets at historic lows, financial constraints imposed by financial institutions under stress directly translated into reduced consumer spending, leading to initially localized, but gradually spreading, cycles of lower corporate sector profitability, higher layoffs, higher unemployment, and more foreclosures.

Household balance sheet vulnerabilities also built up in other advanced economies and several emerging markets. Household debt-to-income ratios rose sharply in several Western European countries (most notably in the United Kingdom, Spain, and Ireland). In several emerging markets, household credit expanded rapidly as well, leading to sharp increases in leverage and vulnerabilities. As real estate prices declined, this adversely affected the quality of loan portfolios and put financial intermediaries at risk, especially in markets where values had risen rapidly.

The large number of individuals involved, the limited information available, and the social repercussions associated with household debt restructuring complicated and slowed down the policy response. While corporate debt restructuring is costly and painful, there are well-established international best practices for confronting widespread corporate defaults. In the case of households, moral hazard problems, the sheer number of cases, and equity and distribution issues complicated the picture. In the United States, notwithstanding political support for a relief package for mortgage holders, policy action on this front was slow and erratic, and no effective solution emerged. In Eastern Europe, several countries were confronted with similar problems, but have yet to respond in a systematic manner.

The Channels and Mechanisms in the Turmoil Stage

The crisis was the first global financial crisis since the Great Depression. Through several phases, its spread was unprecedented in scope and ferocity, with many channels of transmission. It called for large government interventions, which have left many legacies for the future.

The crisis was unprecedented in its spillovers. As in any financial crisis, there are catalysts, triggers, and amplification mechanisms. The catalyst of the crisis was the overextended U.S. housing and mortgage markets. The trigger was the turnaround in U.S. housing prices, in part related to a cycle of monetary policy tightening, with the subprime sector as the main initiator of subsequent turmoil. While the crisis emerged
in the U.S. subprime market, it quickly broadened to the larger housing markets in the United States and spilled over into other U.S. financial markets (for example, other asset-backed securities). Surprising was the degree and speed of global spillovers, which happened in several phases and through various amplification mechanisms (see figure 14 and box 1).

The first phase was through direct exposure. This phase was largely limited to banks with direct exposure to the U.S. market and affected just a few financial markets, sometimes related to liquidity runs (mainly related to excessive funding in wholesale markets). Through direct exposure to subprime-related assets, problems quickly surfaced among European banks, including in Germany (IKB, July 2007) and France (BNP Paribas, August 2007). The U.S. housing market stress also made housing vulnerabilities in several countries apparent, notably in Western Europe, and triggered funding problems in some markets. The United Kingdom saw a bank run on Northern Rock, which had been relatively more reliant on interbank markets rather than on deposits for funds.

In the meantime in the United States, prospects of a deeper housing downturn and rising defaults quickly instigated broader financial turmoil. Worse-than-anticipated credit deterioration in U.S. subprime mortgages prompted surprising multiple-notch downgrades by major rating agencies, which were unable to assess accurately the risks of complex mortgage-related securities and often criticized for being too closely aligned with the issuer. Downgrades led to sharply widening spreads on asset-backed securities and disrupted liquidity in interbank and commercial paper markets. Disruptions were amplified by fundamental uncertainty and
Box 1. Chronology of the Crisis

August 1–17, 2007. German bank IKB is rescued. BNP Paribas halts redemptions on three funds with subprime exposures; European Central Bank (ECB) and other central banks inject overnight liquidity. German bank Sachsen LB receives bailout.


December 12, 2007. U.S. Federal Reserve (Fed), ECB, Swiss National Bank (SNB), and Bank of Canada jointly announce measures to address short-term funding market pressures; Fed establishes Term Auction Facility (TAF).

March 11–16, 2008. JP Morgan acquires Bear Stearns after Fed provides $30 billion in non-recourse funding; Fed creates Term Securities Lending Facility (TSLF) and Primary Dealer Credit Facility (PDCF) to expand liquidity provision to wider group of counterparties.

September 7, 2008. U.S. government sponsored entities Fannie Mae and Freddie Mac are placed into conservatorship.

September 15–16, 2008. U.S. Investment bank Lehman Brothers files for bankruptcy; U.S. authorities step in to rescue insurance firm AIG.


September 29, 2008. Fed increases currency swap lines; euro area governments inject fund into Fortis; Iceland government buys stake in Glitnir bank.

October 7–8, 2008. Coordinated interest rate cuts by major central banks; Fed creates Commercial Paper Funding Facility (CPFF) to purchase asset-backed and unsecuritized commercial paper directly from highly rated issuers. Iceland government takes control of Glitnir and Landsbanki; U.K. government moves to provide capital to U.K. banks and issues debt guarantees.

October 13–14, 2008. Euro area governments announce measures to provide capital to their banks; U.S. government announces Capital Purchase Program (up to $250 billion) under the TARP.


December 4, 2008. Large joint interest rate cuts in Europe: ECB lowers interest rates by 75 basis points, BOE cuts by 100 basis points, and Sweden’s Riskbank cuts by 175 basis points, amid continuing strains in interbank markets.
opacity regarding counterparty risks. As commercial banks decided to absorb (legally separate) vehicles, their balance sheets were strained. Interbank rates spiked, and the issuance of asset-backed commercial paper contracted sharply.

A second phase of international spillovers was transmitted through asset markets. This happened through liquidity shortages, frozen credit markets, and stock market declines, affecting many more markets (U.K. sterling, euro, and Swiss franc). Initial policy responses aimed at addressing liquidity disruptions were large and unprecedented. Major central banks quickly made liquidity available to local commercial banks. While increasingly larger and more flexible—in maturity and especially in scope of collateral accepted—the ability of liquidity injections to calm interbank markets proved short-lived. Furthermore, approaches varied among countries, requiring modifications and rounds of international coordination. Currency swaps between major central banks were also needed to mobilize U.S. dollar funding in overseas markets.

These unprecedented and numerous efforts were unable to remedy the underlying problems that led to a breakdown in market trust and confidence. Unknown viability of institutions, especially affecting internationally active banks, could not be supplanted by central bank liquidity, which increasingly replaced private transactions. The reliability of credit insurance and the integrity of counterparties, particularly in the massive but unregulated market for credit-default swaps, also came into question, notably through the weakening positions of ultimate insurers.

The third phase occurred through large solvency concerns. In October 2008 large solvency concerns affected systemically important global financial institutions, leading to massive sell-offs and risking a financial meltdown. In this phase, liquidity concerns gave way to solvency worries, against the backdrop of highly leveraged financial systems. The buildup of leverage, especially for U.S. investment banks and European commercial banks, made the system vulnerable to a rapid cycle of forced deleveraging and solvency pressures. As financial institutions incurred large losses and wrote down illiquid securities, solvency concerns across markets fueled a process of rapid deleveraging and forced asset sales. Hedge funds—facing financing constraints and redemption pressures—further fueled a rapid unwinding process. This led to further asset price declines, prompting distressed asset sales and rising need for recapitalization and resulting in further loss of confidence.

Compounding the problems, recognition of insolvency problems was delayed, and resolution frameworks proved haphazard in practice. Deficiencies in resolution frameworks in advanced economies allowed problems to intensify. These deficiencies included lack of scope (for example, no coverage of investment banks and insurance corporations), limited coordination (for example, between deposit insurance and lender-of-last-resort facilities), and slow speed (for example, lack of specific frameworks for bank resolution). Weak market confidence and eroded trust required authorities to intervene in some cases, with unprecedented means.

In this phase, global transmission channels were multiple, including through banks and non-bank financial institutions that were rapidly deleveraging. Despite a coordinated cut in policy rates by major central banks and the extension of guarantees in
some countries, market confidence continued to deteriorate, leading to major failures or near-failures. The collapse of Lehman Brothers, with its major interconnections and exposures, shocked market confidence globally. Uncertainties led to deepening turmoil and runs, including on U.S. money market funds, requiring new interventions. Through its substantial exposures in the credit-default swap market, insurance giant AIG nearly collapsed before receiving substantial public rescue funds.

**Large Government Interventions**

The crisis has required governments to intervene in many ways. As asset prices plunged across markets, the risks of cascading institutional failures and financial meltdown prompted actions by authorities across a wide range of advanced countries in mid-October, marking an overdue transition from concerns about liquidity to concerns about solvency (not unlike previous crisis episodes) and the need for more rapid and substantial recapitalization. The principal forms of intervention were (a) liquidity provision through collateralized lending and other schemes, (b) support for short-term wholesale funding markets, (c) (more extensive) guarantees of retail deposits and other liabilities, (d) purchases or exchanges of non-performing or illiquid assets, and (e) capital injections to banks. Furthermore, monetary and fiscal policy responses became even more accommodative in many countries, but this did not stop the decline. Large external financial support from various sources has been necessary for several emerging markets hit by deleveraging.

The amounts of these interventions have been very large. On the basis of money spent and announced commitments, advanced countries have been most affected, while most emerging-market countries have had less need for capital or other forms of financial sector support (see table 1; see also Claessens 2009a; IMF 2009c). Liquidity provided and guarantees extended were large, amounting to double-digit fractions of gross domestic product (GDP) on average for the group of advanced countries. Capital support has been about 2 percent of GDP. Asset purchases to date have been about the same, 2.5 percent of corresponding GDP. Besides the large direct fiscal costs, which are captured by the figures, there are many contingent costs, which are hard to quantify, such as the insurance schemes for assets or increased deposit insurance limits. Indeed, past experiences suggest that amounts will increase further.

These interventions were necessary, but distortionary. The interventions have generally had the desired effects, namely stabilizing financial systems and regaining some measure of confidence. By nature, however, these measures distort, either directly, as they support financial institutions in nonmarket ways, or indirectly, as they can skew and distort resource allocation. A clear example of the (purposely) distortive nature of financial intermediation is the intervention by central banks, notably the U.S. Federal Reserve, in a number of (short-term) markets, either directly (for example, through the purchase of government bonds) or indirectly (for example, through the various liquidity facilities that aim to support specific financial markets, such as the commercial paper market). Another example is the provision of a guarantee scheme for money market funds in the United
States following large outflows after one fund “broke the buck” (its net asset value fell below $1). The guarantee led to deposit outflows at commercial banks, which prompted an increase in deposit insurance coverage.

An example of how guarantees can distort the allocation of risks between financial institutions and the government is Ireland’s extension of guarantees to its largest banks. Prior to the extension of guarantees, the credit-default swap spreads for the large Irish commercial banks were very high. After guarantees, bank credit-default swap spreads declined sharply, while the sovereign spread increased. Measures like these, numerous in many advanced countries today, distort asset prices and financial flows.

Distortions are not just direct, but also indirect and medium term. The indirect distortions affecting the real sector are more difficult to document, but there are many suggestive examples. In many countries, programs have been put in place to support more lending to small and medium enterprises. But large firms have also been targeted for public support. In Japan, for example, in April 2009 Parliament passed a law to allow for the recapitalization of (larger) nonfinancial firms using public funds through the purchase of preferred shares by the (state-owned) Development Bank of Japan. In the United States, France, and Italy, car companies are being (indirectly) supported. In several countries, there are (largely informal) requirements for local lending as part of financial sector support. All of this has, directly and indirectly, affected international competition in various markets, both financial and real (that is, inefficient “zombie” firms may be created, driving out efficient firms).

Furthermore, the increase in direct state ownership and the large indirect role of the state in the financial sector both risk distorting financial intermediation in a deeper and potentially longer-lasting way. The perverse (long-term) consequences of state-owned banks are well documented and, while in most countries the institutional environment should prevent the worst effects, distorted outcomes may still arise. In addition, there are many other (sometimes unintended) consequences of the interventions. In the United States, for example, the caps on remuneration are affecting not only the incentives of financial institutions being supported through public funds but also those of others. These types of rules and, more generally, the larger role of the state can affect the quality of financial intermediation.

Distortions have extended to the international level. Interventions have affected international capital flows and financial intermediation. Liquidity support provided the first manifestation. Actions in the United States initially focused on providing domestic support, even though interbank market prices suggested significant dollar funding pressures for European banks and emerging markets. For mature markets, it took several weeks for central banks to act on these stresses. And, even after ad hoc bilateral swap lines between central banks were set up and their scope was gradually expanded, market prices continued to suggest problems. For emerging markets, the response was slower and the amounts provided were more limited. Liquidity shortages were keenly felt by many emerging markets. Large external financial support from various sources became necessary as emerging markets were hit by the deleveraging process, but the real consequences had already been incurred.
## TABLE 1. Financial Sector Support and Up-front Government Financing, by Region and Select Countries, as of April 15, 2009

% of 2008 GDP

<table>
<thead>
<tr>
<th>Region</th>
<th>Capital injection (A)</th>
<th>Purchase of assets and lending by the treasury (B)</th>
<th>Central bank support provided with treasury backing (C)</th>
<th>Liquidity provision and other support by central bank&lt;sup&gt;a&lt;/sup&gt; (D)</th>
<th>Guarantees&lt;sup&gt;b&lt;/sup&gt; (E)</th>
<th>Total (A+B+C+D+E)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advanced North America</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>0.0</td>
<td>8.8</td>
<td>0.0</td>
<td>1.6</td>
<td>13.4</td>
<td>23.7</td>
</tr>
<tr>
<td>United States</td>
<td>3.9</td>
<td>1.3</td>
<td>1.1</td>
<td>42.1</td>
<td>31.3</td>
<td>79.6</td>
</tr>
<tr>
<td><strong>Advanced Europe</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>5.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>30.0</td>
<td>35.3</td>
</tr>
<tr>
<td>Belgium</td>
<td>4.7</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>26.2</td>
<td>30.9</td>
</tr>
<tr>
<td>France</td>
<td>1.2</td>
<td>1.3</td>
<td>0.0</td>
<td>0.0</td>
<td>16.4</td>
<td>19.0</td>
</tr>
<tr>
<td>Germany</td>
<td>3.8</td>
<td>0.4</td>
<td>0.0</td>
<td>0.0</td>
<td>18.0</td>
<td>22.2</td>
</tr>
<tr>
<td>Greece</td>
<td>2.1</td>
<td>3.3</td>
<td>0.0</td>
<td>0.0</td>
<td>6.2</td>
<td>11.6</td>
</tr>
<tr>
<td>Ireland</td>
<td>5.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>257.0</td>
<td>263.0</td>
</tr>
<tr>
<td>Italy</td>
<td>1.3</td>
<td>0.0</td>
<td>0.0</td>
<td>2.5</td>
<td>0.0</td>
<td>3.8</td>
</tr>
<tr>
<td>Netherlands</td>
<td>3.4</td>
<td>2.8</td>
<td>0.0</td>
<td>0.0</td>
<td>33.7</td>
<td>39.8</td>
</tr>
<tr>
<td>Norway</td>
<td>2.0</td>
<td>15.8</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>17.8</td>
</tr>
<tr>
<td>Portugal</td>
<td>2.4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>12.0</td>
<td>14.4</td>
</tr>
<tr>
<td>Spain</td>
<td>0.0</td>
<td>4.6</td>
<td>0.0</td>
<td>0.0</td>
<td>18.3</td>
<td>22.8</td>
</tr>
<tr>
<td>Sweden</td>
<td>2.1</td>
<td>5.3</td>
<td>0.0</td>
<td>15.3</td>
<td>47.3</td>
<td>70.0</td>
</tr>
<tr>
<td>Switzerland</td>
<td>1.1</td>
<td>0.0</td>
<td>0.0</td>
<td>10.9</td>
<td>0.0</td>
<td>12.1</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>3.9</td>
<td>13.8</td>
<td>12.9</td>
<td>0.0</td>
<td>51.2</td>
<td>81.8</td>
</tr>
</tbody>
</table>
### Advanced Asia and Pacific

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Japan</strong></td>
<td>2.4</td>
<td>11.3</td>
<td>0.0</td>
<td>1.2</td>
<td>7.3</td>
<td>22.1</td>
<td>0.8</td>
</tr>
<tr>
<td><strong>Korea, Rep. of</strong></td>
<td>2.7</td>
<td>5.4</td>
<td>0.0</td>
<td>0.3</td>
<td>13.8</td>
<td>22.2</td>
<td>0.4</td>
</tr>
</tbody>
</table>

### Emerging Economies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Argentina</strong></td>
<td>0.0</td>
<td>0.9</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.9</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>Brazil</strong></td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.5</td>
<td>0.0</td>
<td>1.5</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>China</strong></td>
<td>0.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.5</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>India</strong></td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>5.6</td>
<td>0.0</td>
<td>5.6</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>Indonesia(^n)</strong></td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td><strong>Hungary</strong></td>
<td>1.1</td>
<td>0.0</td>
<td>0.0</td>
<td>4.7</td>
<td>1.1</td>
<td>6.9</td>
<td>1.1</td>
</tr>
<tr>
<td><strong>Poland</strong></td>
<td>0.4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.1</td>
<td>3.6</td>
<td>0.4</td>
</tr>
<tr>
<td><strong>Russian Federation</strong></td>
<td>0.3</td>
<td>0.5</td>
<td>3.2</td>
<td>3.2</td>
<td>0.5</td>
<td>7.7</td>
<td>0.8</td>
</tr>
<tr>
<td><strong>Saudi Arabia</strong></td>
<td>0.6</td>
<td>0.6</td>
<td>0.0</td>
<td>8.2</td>
<td>–</td>
<td>9.4</td>
<td>1.2</td>
</tr>
<tr>
<td><strong>Turkey</strong></td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.2</td>
<td>0.0</td>
<td>0.2</td>
<td>0.0</td>
</tr>
</tbody>
</table>

### Average (PPP GDP weights)

<table>
<thead>
<tr>
<th></th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
<th>2012</th>
<th>2013</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>G-20</strong></td>
<td>1.9</td>
<td>2.5</td>
<td>1.0</td>
<td>12.4</td>
<td>14.3</td>
<td>32.1</td>
<td>3.4</td>
</tr>
<tr>
<td><strong>G-20 European Union</strong></td>
<td>2.7</td>
<td>3.8</td>
<td>3.2</td>
<td>0.5</td>
<td>22.1</td>
<td>32.3</td>
<td>6.7</td>
</tr>
<tr>
<td><strong>Advanced economies</strong></td>
<td>2.9</td>
<td>4.0</td>
<td>1.3</td>
<td>18.8</td>
<td>22.8</td>
<td>49.8</td>
<td>5.3</td>
</tr>
<tr>
<td><strong>Emerging economies</strong></td>
<td>0.2</td>
<td>0.1</td>
<td>0.4</td>
<td>1.6</td>
<td>0.1</td>
<td>2.4</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Sources: IMF 2009d; IMF, Fiscal Affairs Department Monetary and Capital Markets database on public interventions.

Table continues with notes on p. 262.
Guarantees on deposits and other liabilities issued by individual countries have led to “beggar-thy-neighbor” effects, forcing other countries, starting after Ireland, to follow with similar measures. Some advanced countries, especially those closely integrated (such as the European Union and the European Monetary Union) quickly coordinated policies, such as the adoption of uniform deposit guarantee coverage. The rapid spread of guarantees led to further financial turmoil in other markets. Many emerging markets, not able to match the guarantees, suffered from capital outflows as depositors and other creditors sought safe havens. Distribution of risks sharply changed over time and across circumstances. Furthermore, policy measures to encourage lending often were biased toward local lending, putting international operations at a disadvantage.

Countries were quick to “ring fence” assets in their jurisdictions when cross-border entities showed signs of failing, reflecting the absence of clear burden-sharing mechanisms for banks with international operations. Examples of defensive “asset grabs” were multiple. One was the decision by U.K. supervisors, fearing an imminent collapse of Icelandic bank branches (under the authority of Icelandic supervisors, who did not commit to recompense U.K. bank liabilities), to resort to the Anti-terrorism, Crime, and Security Act to ring fence Icelandic bank assets within the United Kingdom. Another one was the German initiative to freeze Lehman’s assets to assure the
availability of cash to satisfy depositors before they could be claimed by the parent under U.S. bankruptcy proceedings. Such actions constituted anticompetitive behavior in that they tended to favor local interests.

Few actions were internationally coordinated. Most government interventions to date have been at national levels. Although there were some coordinated actions (for example, among Belgium, the Netherlands, and Luxembourg, with some involvement of France, to resolve Dexia and Fortis), these were driven largely by national interests (as suggested by the fact that the intervened entities were often broken up along national markets and in line with support). The main exception was the coordinated (after some serious disruptions) provision of liquidity support. In the euro area, central bank actions are, by design, (nearly) fully coordinated among euro system members.

**The Current State of Affairs and the Need to Plan for Exit**

The crisis is still evolving. The financial turmoil and rapid economic slowdown in advanced countries continue to affect global markets. This has happened through both financial (cross-border banking, hedge funds) and real economic channels. Starting in late 2008 and intensifying in 2009, the drop in demand in major advanced countries affected many markets, with sharp drops in exports in many emerging markets. With recessions and economic slowdown affecting all countries, the scope for export-led growth sharply diminished, depriving especially those countries with large foreign exchange exposures of a potential channel of recovery. These recessions had adverse effects on financial sectors around the world, raising nonperforming loans and further weakening capital adequacy positions. Cross-border exposures were (and are) large factors behind the weakening of banks in many markets.

Continued turmoil means that extraordinary government interventions will continue, and the (international) rules of the game will remain in flux. The coverage and scope of interventions and other policy measures will evolve depending on their effectiveness, and the conditions and the amount of support will likely increase further. As circumstances evolve, governments will (need to) adjust the rules, such as how to treat shareholders and creditors when restructuring large financial institutions, and this will create further uncertainty. If political support diminishes, support for financial sector restructuring may become (even) more nationally oriented, and distortions will increase further.

Governments need to plan for exit. While serious risks remain, calling for more interventions, it is generally agreed that distortions should be removed as quickly as possible to return to a sustainable system in line with a new financial architecture. As the crisis abates, governments need to plan for exit from guarantees, large deposit insurance, ownership, asset acquisitions, and so forth. They have to do this within their fiscal constraints. These are difficult processes, and many are unprecedented, especially in the context of highly integrated financial systems. They will require some coordination. It is clear, however, that lack of coordination can create (new) distortions. If the unwinding of interventions is not coordinated internationally, it can
aggravate still weak confidence, create new distortions, and potentially be anticompetitive. Especially for the removal of guarantees, governments would do well to coordinate to avoid large capital movements. Yet, while desirable, more coordination will be difficult in practice.

Lessons for Macroeconomic Policy and for National and International Financial Reforms

The crisis has reopened the debate on whether economic policy should be concerned with asset price booms and increases in leverage. It has highlighted, in abundantly clear ways, the deficiencies in national financial regulation and supervision in many advanced and some emerging markets and developing countries. And it has highlighted how the international financial architecture has fallen behind a rapidly integrating international financial system. These are broad reform agendas for the future, which deserve more than the summary of thinking provided here.

Lessons for Macroeconomic Policy

What should be done regarding asset price booms and increases in leverage? Should economic policy be concerned with financial market developments? Should policy be used to dampen booms? And, if so, does this fall under the responsibility of monetary policy? What, if any, should be the role of fiscal policy? This debate has been going on for some time and will continue to occupy economists and policy makers for a while. On the basis of previous research, however, and with the current financial crisis as another input, a few preliminary conclusions can be reached (for a more detailed review, see IMF 2009c).

For one, not all booms are alike. What may matter is not so much the asset price boom in itself, but who holds the assets and the risk, how the boom is financed, and how an eventual bust may affect financial institutions. The degree of leverage associated with the funding of a boom and the degree of involvement of banks and other financial intermediaries will determine the magnitude of balance sheet effects and the dangers to the supply of credit in a bust.

It is also likely that the case for policy intervention depends on how a boom is financed and how risk is held. Asset price booms supported through leveraged financing and involving financial intermediaries should be dealt with, since they entail risks for the supply of credit to the economy; other booms could more likely be left to themselves. This latest boom, financed by banks and through the shadow banking systems, has been much more costly than the Internet bubble of the late 1990s, which was financed largely by equity markets.

One lesson given the risks of a leveraged boom is that the mandate of monetary policy should include macrofinancial stability, not just price stability. To the extent that the buildup of systemic risk can portend a sharp economic downturn, and to the extent that regulation cannot fully prevent such a buildup, it is now clear that
policy makers cannot neglect asset price and credit booms. That said, the crisis also confirms that prudential measures provide a more targeted and less costly policy solution than interest rate changes and should be a central element of an integrated policy response.

**Lessons for Fiscal Policy**

The crisis also highlights two important lessons for fiscal policy. The first is that in many countries, budget deficits were not reduced sufficiently during the boom years when revenues were high, which limits the fiscal space needed to fight the crisis. The second has to do with the structure of taxation. In most countries, the tax system is biased toward debt financing through deductibility of interest payments. This bias toward higher leverage increases the vulnerability of the private sector to shocks and should be eliminated.

**Lessons for National Financial Architecture**

Regulatory shortcomings have clearly contributed to the financial crisis (for a more detailed review of needed national financial architecture reforms, see IMF 2009b). The recognition of these failures is driving the current redesign of regulatory and supervisory systems across a large range of countries. Coordinated by the Financial Stability Board (previously named the Financial Stability Forum), national authorities and standard setters are working to address deficiencies in the existing arrangements. This is a broad agenda that will continue for some time. A summary of overall objectives and current thinking is nevertheless useful.

Actions are generally recognized as required in five general areas:

- **Regulatory perimeter.** The regulatory, supervisory, and information perimeter needs to be broadened to ensure that all financial activities that pose systemic risks are adequately captured.

- **Microprudential regulation.** Capital regulation, liquidity management, and risk management need to reflect not only the risks of individual institutions but also their potential to form systemic risk.

- **Macroprudential regulation.** Regulatory approaches need to be designed that better dampen the procyclicality of financial markets.

- **Information and market discipline.** Improved information disclosure and corporate governance practices are needed to enhance market discipline.

- **Organization of regulation and supervision.** Greater coordination is needed within and across countries in both the design of regulation and the monitoring of systemic risk.

The following key principles are recognized as essential guides to these redesigns. First, the perimeter of regulatory and supervisory arrangements should be drawn to address concerns over systemic risk and be compatible across jurisdictions, institutions,
and activities. This means that supervisory authorities need to identify and address gaps in oversight and information since markets and institutions will otherwise seek to exploit them. In that context, supervisory resources should be increased and allocated to the areas posing greatest systemic risk. Supervisory actions should result in prompt intervention whenever excessive risks arise.

Second, regulations need to be incentive compatible while balancing possible adverse impacts on innovation and efficiency. This means that regulation should provide any institution whose distress would have systemic externalities with incentives to internalize such costs in its business planning and risk management.

Third, market discipline and supervision should complement each other. This means allowing for the failure of individual institutions. This should occur within the context of a credible resolution framework for banks and non-bank financial institutions that limits the wider impact of failure and reduces the moral hazard of a bloated public safety net. It also requires improved corporate governance and information disclosure.

Finally, the redesign of financial regulation needs to be aware of and seek to overcome its inherent limitations. Many questions remain about how to best reform the architecture to mitigate systemic risks effectively without imposing too much and inefficient regulation. Many recent rules are still in the process of being implemented. The redesign needs to keep regulatory burdens in mind. At the same time, regulation tends to lag behind financial innovation and is vulnerable to industry capture and political influence. Supervisors may lack the mandate, resources, or independence to contain systemic risk, and enforcement may be poor. These limits on regulation and supervision are especially binding in emerging markets and developing countries. Implementing the new rules will thus remain especially difficult in those countries.

Lessons for International Financial Architecture

Many changes in international financial architecture are needed, including regarding surveillance (for a more detailed review of needed financial reforms, see IMF 2009c). The crisis has revealed the enormous costs of not identifying risks early enough. Private market discipline failed in many respects, while public surveillance identified risks at a broad level but did not drill deep enough to expose the full extent of vulnerabilities or draw specific policy conclusions. Many changes are needed to reduce systemic risks globally. A more effective approach to detect impending dangers to the world economy will require close cooperation among international agencies to bring together the scatter of macrofinancial information and expertise and to identify key risks and vulnerabilities. Only by working across organizations—supported by significant information sharing and drilling down—can one hope to “connect the dots” (across financial institutions, markets, and countries), clearly articulate risks, and propose practical remedies.

Obtaining better information is another essential step. More, and better organized, information is required for markets and policy makers to improve the assessment of systemic risk. The crisis has underlined the importance of going beyond traditional statistical approaches to obtain timely and higher-frequency real and financial indicators,
at least for systemically important countries and financial institutions. This requires enhancing the accessibility and timeliness of existing data, developing new sources of information, and promoting transparency and disclosure more generally. Data need to cover non-bank financial institutions, such as insurance companies and hedge funds, and housing-related statistics, which would allow a better understanding of credit risk transfers. Better information is needed on the financial operations of large nonfinancial corporations that have significant links in national economies and potentially across borders as well.

Better risk assessment means strengthening macrofinancial analysis and work on early-warning systems. More analysis is needed on the linkages between financial sector and macroeconomic performance (for instance, on the relationship between monetary policy and financial risk taking). And new and better operational tools need to be developed for macrofinancial surveillance. Perhaps most critical is recognizing that early-warning exercises are less about “calling” crises—whose exact timing and occurrence are nearly impossible to foretell—than about identifying risks and underlying vulnerabilities and taking remedial policy actions. But even then, it is possible to miss new channels through which identified risks and novel manifestations of risk can spread, especially as financial innovation and integration continue and the complex web of interlinkages grows.

Early-warning and surveillance work by multilateral agencies will need to balance voluntary engagement in assessments with mandatory compliance. Multilateral and bilateral assessments could be used more systematically to examine macroprudential risks and implement multilaterally agreed principles, standards, and actions. It will, however, mean stronger requirements on member regulators and authorities to participate, more streamlined processes, and improved means of dissemination, while recognizing the tension inherent in the function of whistleblower and crisis preventer. More broadly, an overarching challenge in improving early warning will be to convince country authorities to take actions to deal with vulnerabilities, particularly during good times. Change in international financial governance and representations (in both rule-making and decision-making bodies, such as the Financial Stability Board, the Basel Committee on Banking Supervision, the International Monetary Fund, the G-7, the G-9, and the G-20) will be needed to make this effective.

Better cross-border crisis management arrangements are also needed (see Claessens 2009b). As clearly demonstrated by the failures of Lehman Brothers, AIG, and some Icelandic banks, countries cannot deal with large, complex, globally active financial institutions on their own, as these institutions affect many markets and countries. A more universal approach is needed. Closer cooperation and greater coordination among regulators and supervisors could help to address market disruptions as they arise and forestall policy measures that have adverse spillovers. An enhanced role for “colleges of supervisors” with specific mandates and accountability will be an important component to achieve the goal of better monitoring and early interventions. At the same time, this will not be sufficient to cover all sources of systemic risks, as risks can come from other sources, including non-bank financial institutions.
Cross-border banking resolution needs to be improved. Clear and binding rules on burden sharing for weak or failed cross-border financial institutions are needed; otherwise it will be hard to develop a fail-proof system. The best system—a global financial regulator, matching the current, financially closely integrated world and well resourced in staff, powers, budget, and financial resources—is unlikely to materialize soon. Other options, each of which could achieve varying degrees of global financial stability, are a new charter for internationally active banks, greater harmonization of rules and practices, and enhanced coordination. Each of these second-best reforms has its own benefits and costs, which are difficult to rank, especially as they depend on implementation and enforcement.

Better liquidity provision is needed. Improved crisis management will require better provision of international liquidity, to both financial institutions and countries, to prevent spillovers from becoming solvency issues. While the designs and institutional frameworks can build on those for national lender-of-last-resort facilities, much work is needed to obtain better facilities for cross-border banks. Many of the obstacles are similar or relate to the same underlying factors hindering ex post crisis resolution. For liquidity provision at the country level, the approaches are conceptually also well known and can involve, besides private market solutions (including contingent credit lines and insurance contracts), bilateral or regional swaps among countries, other forms of reserve pooling, and an expanded International Monetary Fund, including a larger allocation of special drawing rights (as agreed in principle at the G-20 meeting of April 2, 2009). But many barriers may exist between principles and practices.

**Short-Term Issues and Lessons for Emerging Markets and Developing Countries**

The crisis has affected emerging markets and developing countries in many ways through financial and real economic channels. Countries have had to respond, often in new and unorthodox ways, and many still face challenges on multiple fronts. Policy choices are greater today than in the past, since many countries have entered the crisis with better fiscal positions, stronger financial sectors, and improved institutional frameworks. The crisis also provides some lessons on the medium-term financial sector development strategies for emerging markets and developing countries and on how to adapt policies and reforms to their specific circumstances.

**Short-Term Issues and Policy Options**

The current financial turmoil is confronting emerging-market economies and developing countries with two shocks: a “sudden stop” of capital inflows driven by global deleveraging (see figure 15) and a collapse in export demand associated with the global slump in trade (see figure 16 and the summary of Ghosh and others 2009). Although some countries were ripe for a homegrown crisis following unsustainable credit booms or fiscal policies and are facing large debt overhangs, the majority were innocent bystanders (see figure 17). Policies to address the current situation and bring about recovery in both groups of countries involve obtaining more external financing, adjusting monetary and
fiscal policies, and being prepared to address the many debt restructurings, all while considering the exit strategy from what are often unorthodox policies.

A key ingredient to address these two shocks will be greater official financing. Faced with massively lower inflows, and sometimes outflows, of private capital (see figure 18), countries need official external financing to expand their “policy space.” Such funds could allow emerging markets and developing countries to pursue supportive macroeconomic policies—including, in countries with large debt overhangs, by helping to meet the fiscal outlays (such as costs to recapitalize banks) associated with resolving those overhangs. Another key ingredient will be policies to protect the poor and other vulnerable groups.

For their short-term external financing needs, some countries can seek recourse to swap facilities from major advanced-economy central banks. International financial institutions, including the IMF through its new and existing instruments, can provide

**FIGURE 15.**
Net Inflows in 2007 and Net Outflows in 2009 for Emerging Economies

a. Net flows to emerging economies (fourth quarter)

![Net flows to emerging economies](image)

b. Emerging-market sovereign and corporate external spreads

![External spreads](image)
some such support as required, both through direct balance-of-payments support and in a contingent fashion through credit lines. Some countries will be able to use lines for trade credit made available by bilateral and multilateral financial institutions, although it is not clear that there are large market failures in the provision of trade finance (see Chauffour and Farole 2009).

Monetary policy can be eased and, for many countries, the exchange rate can adjust. Except where the loss of confidence in the currency precludes it, the basic thrust of monetary policy should be toward easing. The global deflationary pressures and widening interest differentials with respect to advanced countries allow much room for lower interest rates. Quantitative measures may also be appropriate in some cases. However, central banks need to remain mindful of the trade-off between the growth-enhancing effects of looser policy and the negative impact of exchange rate depreciation on unhedged balance sheets.
Countries with flexible exchange rate regimes should allow their exchange rate to absorb much of the pressures (see figure 19). Since many emerging markets have large stocks of foreign exchange reserves, they can use some of these for intervention to avoid disorderly market conditions. Using reserves can prevent excessive depreciation and smooth the impact of interruptions or reversals in capital flows, recognizing that some changes in capital flows may be permanent. In some cases, foreign exchange reserves can also be used to substitute for foreign credit lines to banks, allowing the latter to maintain domestic lending operations. For some whose currency is pegged to the exchange rate, there may be scope to increase the flexibility of the exchange rate regime, while maintaining a credible anchor for monetary policy.

Depending on the available fiscal space, expansionary fiscal policy could also be deployed to support economic activity. With a better fiscal position, many countries can not only allow automatic stabilizers to operate, but also increase discretionary spending (see figure 20). Although the empirical evidence is not conclusive, conventional fiscal multipliers may be relatively small in emerging markets and developing countries, and the impact of fiscal stimulus on activity is more uncertain. This calls for a variety of fiscal measures that could include some less conventional steps such as providing credit guarantees. For countries in crisis, options are more limited, given the potential adverse feedback between debt sustainability and (real) domestic interest rates (see figure 21). For these countries especially, fiscal support should be geared toward maintaining financial sector confidence and solvency.

An important policy step is to prepare for financial turmoil and insolvencies. Governments need to have contingency plans for limiting the risks of bank runs and adopt comprehensive mechanisms to reduce the risks of systemic solvency problems. Many countries need to ensure an adequate framework to facilitate rapid debt workouts. Debt restructuring mechanisms can provide greater scope for monetary easing by reducing
the negative repercussions of exchange rate depreciation on unhedged balance sheets. Depending on circumstances, restructuring can be done either ex post (recapitalizing banks after they suffer losses) or more proactively. However, the large outlays required to restore banks’ solvency may limit the room for conventional fiscal expansion.

It is critical that countries have a credible exit strategy. Monetary policy should not be loosened too quickly, as a rapid reversal would damage credibility. The same holds for fiscal policy interventions, where the stimulus should not be withdrawn too soon. At the same time, stimulus may require a credible exit strategy that places government finances on a long-term sustainable footing. This would help to contain the costs of financing the short-term stimulus and have an additional benefit of strengthening investor confidence and facilitating the resumption of capital inflows in the recovery phase.
Conclusions

The financial crisis has brought a number of weaknesses in economic policy and national and international financial architecture into the open. The reform agenda is large, much remains to be done, and new questions have come up for the design of national and international financial systems. Much has been achieved, but the financial sectors in emerging markets and developing countries still face specific challenges.

Notes

1. This section is based on joint work with Giovanni Dell’Ariccia, Deniz Igan, Luc Laeven, and Krishna Srinivasan.


References


Stijn Claessens’s paper offers a comprehensive review of the causes of the current financial crisis in comparison with past crises and makes many sensible recommendations regarding the national and international financial architecture. It is, without a doubt, an excellent paper.

His paper concludes, “The financial crisis has brought a number of weaknesses in economic policy and national and international financial architecture into the open. The reform agenda is large, much remains to be done, and new questions have come up for the design of national and international financial systems.” I could not agree more.

In that connection, let me focus on three issues, which in my view could have been explored further in Claessens’s paper. Those are institutional aspects of national financial architecture, the international monetary system, and global imbalances.

**National Financial Architecture**

An important question to ask is, what lessons can we draw from the current crisis? There is a danger that emerging markets and developing countries may learn the wrong lessons from the current crisis. What appear to be the “right” lessons for advanced markets may not always be the right ones for developing countries. For the advanced markets, for example, stronger or expanded supervision of those outside of regulatory oversight and reregulation may be the right lessons. However, for many emerging markets and developing countries, the opposite—that is, enhanced deregulation of their financial system—appears to be the key to the problems they face today. Many structured products and derivatives, of which opaqueness and complexity contributed to the buildup of systemic risk in the advanced markets,
have not been widely introduced in most emerging markets. The emerging markets’
direct exposure to these products has been very small. The problems in emerging
financial markets are still too much regulation, too much control, and too many
interventions by the government, stifling innovation and efficiency improvement in
the financial sector. The current crisis should not give reasons for the governments
or politicians to overreact by further strengthening government control and inter-
ventions. In most emerging markets, it is not too little regulation that has to be
addressed, but too much of it. As the title of Gary Gorton’s recent paper, “Slapped
in the Face by the Invisible Hand” (Gorton 2009) suggests, too much reliance on the
invisible hand may be guarded at all times in advanced economies. However, in
developing economies and emerging markets, the invisible hand is still tied behind
the back or has not become strong enough to slap the face of the market. If the les-
sions learned from the crisis push developing countries to move toward stronger reg-
ulation, I fear that the future innovations and development of their financial systems
will face greater challenges.

Having said that, I believe that the current crisis raises questions common to both
advanced and emerging economies in relation to the shape of the financial industry.
Those are, first, would it be wise to encourage, through the regulatory framework,
the emergence of mega banks and financial conglomerates? And, second, how can we
strengthen systemic regulation and macroprudential regulation?

Regarding the first question, many emerging markets have followed the trend of
regulatory reforms in the international money centers during the last decade or two.
Now they may have to rethink their policies promoting mega banks, financial con-
glomeration, and a universal banking system. As Claessens discusses in his paper, one
of the causes of the current crisis is the increased interconnectedness among financial
institutions and markets. In the current financial market environment, it is true that
the banks and other financial institutions have to be big enough to compete in inter-
national markets. In order for banks in emerging markets to be internationally com-
petitive, they have to be much bigger relative to the size of their economy than banks
in advanced economies. Considering the systemic importance that these mega banks
will have, the government is likely to become the captive of these large banks and
financial conglomerates. If all banks become too big to fail and government cannot
let them exit from the market even in the case of mismanagement, then this is no
longer a sound market economy. Economies of scale and scope are important, but
increased interconnectedness can increase the systemic threat of a mismanaged indi-
vidual institution. The current crisis poses a question for both emerging and
advanced markets regarding the architecture of the financial industry: to what extent
should the government encourage the conglomeration of the financial industry and
mergers among banks and non-bank financial institutions? There is an obvious trade-
off between competitiveness of individual financial institutions or groups and long-
term stability of the financial system as a whole.

Regarding the second question, it has become clear that the supervision of indi-
vidual institutions would not have been sufficient to prevent the crisis. Systemic reg-
ulation is also required—that is, regulation that not only takes into account an indi-
vidual institution’s risk but also its potential to form systemic risk. At the same time, the importance of macroprudential regulation must be highlighted. In that connection, the roles and responsibilities of the central bank (monetary authority), regulatory bodies, and finance ministry (financial sector policy or architecture) have to be reviewed carefully to confirm whether each of them has adequate power to deal effectively with a crisis. If necessary, their roles need to be redefined. The recent U.S. proposal to strengthen the role and responsibility of the Federal Reserve Board will provide momentum for many other countries to rethink the role of their own institutions. Given its responsibility to provide liquidity support in times of trouble, the central bank needs to have firsthand information on the status of financial institutions that could pose a systemic risk.

**International Financial Architecture**

In his paper, Claessens says that “international financial architecture is still far from institutionally matching the closely integrated financial systems.” I absolutely agree. This is one of the fundamental causes of the current global financial crisis. The global financial market has been integrated like a single market, but there is no single authority or institution that governs the supply of liquidity and oversees the global financial market. This asymmetry poses great challenges.

Let me focus mainly on the problems faced by emerging and developing economies in this regard. Most emerging-market currencies are not convertible. With global financial integration comes an increasing proportion of the assets and liabilities of banks and non-bank financial institutions that is denominated in international settlement currency, especially the U.S. dollar. This leaves national governments and central banks in a helpless position when facing a dollar liquidity problem. For example, during the period of deleveraging in the United States and Europe, emerging-market financial institutions faced severely constrained liquidity. The central banks of international settlement (or reserve) currency, including the U.S. Federal Reserve Board, the European Central Bank, and the Bank of England, provided liquidity to their domestic institutions, but they did not, and could not, do the same to foreign institutions, despite the fact that both domestic and foreign institutions made transactions in the same currency. This deepened the instability of the exchange and financial markets of emerging-market economies. The balance sheet effect of exchange rate volatility further deepened the crisis in these economies.

As Claessens says in his paper, many emerging economies have a short history of financial market opening, but they are increasingly more integrated into international financial markets. Advanced countries have gradually opened their financial markets during the last 50–60 years, whereas most emerging-market economies only began to open in the 1990s. However, with regard to the size of assets and capital flows relative to the size of the economy, many emerging-market economies are more integrated with the international markets than advanced economies (see figure 1).
Regarding the share of foreign banks in the banking industry or foreign investors in domestic stock markets, emerging markets are more heavily exposed to international transactions than advanced markets. Yet their institutions and systems have not been established to an extent that will allow them to weather the problems associated with high exposure to international capital flows. The root of this challenge goes back to the nonconvertible nature of their currencies. As such, emerging-market economies are more vulnerable to the contamination and spillovers of crises originating elsewhere.

Countries have two options to consider in this case. First, they can reverse the process of integration or at least reduce the pace of integration. Second, they can build up a strong safety net against a crisis induced by increased integration. In fact, the second option might be the only practical one, since they would find it difficult to give up the advantages generated by greater integration. What would such a safety net entail?

1. Accumulation of large amounts of foreign reserves
2. Readily available foreign liquidity support by international financial institutions such as the International Monetary Fund
3. Liquidity support by central banks of reserve currency countries

With regard to the first option—the accumulation of foreign reserves—before the current crisis, concerns had been expressed regarding the buildup of large foreign reserves, mainly because large reserves were seen as too costly and too distortionary for the exchange rate. If, however, the Republic of Korea had not accumulated such a large amount of reserves, it might have become the victim of another currency crisis. The costly process of building up reserves paid off to a large extent. Korea had foreign reserves amounting to $270 billion, ranking fifth or sixth largest in the world, and yet Korean authorities were not able to assuage anxious foreign investors and lenders. If every emerging economy would try to hold such large amounts of foreign reserves as insurance against crisis, the consequences would be damaging, including

![Figure 1](image_url)

**FIGURE 1.**
Share of Foreign Investors in the Stock Market in Select Economies, 2006

Source: Korea Financial Supervisory Service.
mercantilistic domestic policies, undervaluation of the currency, promotion of exports, and continuation of global imbalances.

Regarding the option of foreign liquidity support, International Monetary Fund resources were too small, too slow, and too costly (in terms of policy conditionality) to be reliable. Korea again regarded such support as the last option, given its experience during the 1997 crisis.

In my view, the third option—liquidity support by central banks of reserve currencies—has to be explored further. A global central bank is not going to be established any time soon, so the role of the central bank of reserve currency countries has to be extended beyond the national border. According to recent Korean experience, a swap arrangement with the Federal Reserve Board is the most effective way to stabilize the exchange market. Even such a large amount of foreign reserves failed to quiet the worries and rumors of a shortage of dollars. The swap arrangement between the Federal Reserve Board and the Bank of Korea, however, was quite successful. A global central bank would be the first-best solution. But extending the role of lender of last resort of the central banks of reserve currency countries beyond their national boundaries could be the second-best approach. According to Claessens, “The best system—a global financial regulator, matching the current, financially closely integrated world and well resourced in staff, powers, budget, and financial resources—is unlikely to materialize soon. Other options, each of which could achieve varying degrees of global financial stability, are a new charter for internationally active banks, greater harmonization of rules and practices, and enhanced coordination.” I believe that the same suggestion could be applied to the monetary authority. Swap arrangements between the central bank of a reserve currency country and the central banks of non-reserve currency countries should be expanded and strengthened. Of course, moral hazard problems can arise, but they could be addressed through the Financial Stability Board or through the International Monetary Fund’s surveillance.

A regional monetary arrangement is another option to consider. Swap arrangements among central banks in the region or a regional monetary fund could also help to protect participating countries from the risk of a currency crisis.

**International Standards and Rule Making**

Again, I agree with Claessens on the role of developing countries in international rule making. Many of the international standards currently in place have a bias toward the circumstances of advanced countries. The question remains whether it is legitimate and also relevant to apply those standards to every country in the world. For instance, the currencies of developing countries are not international settlement currencies. Thus a large part of the transactions of their companies and banks is denominated in international currencies. According to current global accounting standards, the books of firms and banks in developing countries have to reflect the current exchange rate. When the exchange market becomes unstable, the debt ratio, capital adequacy ratio, and other financial indicators of firms and banks in these economies
become volatile. This, in turn, can magnify the instability of the financial market and the real economy.

To date, developing countries’ participation in global forums such as the Financial Stability Forum, the Basel Committee on Banking Supervision, and other groups has been limited. The representation of developing countries and emerging markets in these groups will have to be enhanced. I fully support Claessens’s argument that some overrepresentation in standard-setting bodies and tilting of the bargaining position toward developing countries would be helpful.

Global Imbalance

Claessens’s paper does not address the issue of global imbalances at all. As many economists agree, the current crisis has much to do with long-sustained global imbalances. Unless we tackle this issue, the discussion of the sustainable global economic system cannot be complete. The key issue here is to share responsibilities among major economies. The most important economies in the global economy are the United States, Europe, and China. Among these three, the Chinese currency is not yet convertible.

The savings and investment gap between the United States and China has to be addressed. In this regard, China will have to expand its domestic consumption and corporate investments through reforms in pension, housing, health care, capital markets, and corporate governance. At the same time, the renminbi should be made convertible and allowed to float more freely in the market. The exchange control and capital market restrictions should be reduced. This would lead to more proper valuation of the Chinese currency and to the rebalancing of current account positions among these countries. Ultimately, the big discrepancy in the exchange rate regime among major economies will have to be addressed if we are to correct the global imbalances and achieve sustainable global economic growth.

Reference

A Sustainable Global Economic System after the “Great Recession”? Some Lessons from History

GIOVANNI ZANALDA

The “Great Recession” has sparked a renewed interest among scholars and policy makers in the study of past financial crises (generally divided into banking, currency, debt, and combinations of the three), their origin, unfolding, and consequences. Historical precedents show that lasting solutions are reached by having a clear understanding of both “monetary” and “real” determinants of a crisis as well as of its international implications. Crises then become an opportunity to implement reforms at the domestic and international levels to put the global economy onto a sustainable path of growth. In the immediate aftermath of the 2008 crisis, governments around the world have been collaborating to prevent a complete meltdown of the global financial system. G-20 meetings, increased reliance on (and funds to) the International Monetary Fund, implementation of large fiscal stimuli, and expansionary monetary policies have, for the time being, signaled a willingness to stabilize the world economy and prevent a rush to “beggar-thy-neighbor” types of policies. However, history shows that this encouraging trend can be readily reversed if countries are not able to maintain expansionary policies while implementing reforms at both domestic and global levels.

The fall of Lehman Brothers in September 2008 has changed our perception of the workings of global finance and of the stability and sustainability of the type of growth the world has experienced since the 1980s. First, the growth of the world economy in the last two decades was based on vast global current account imbalances. On the one hand, the United States accounted for most of the deficit, and on the other, Japan, Germany, China, and other emerging economies accounted for most of the surplus. Second, the international expansion of financial and credit markets was rooted in the assumption that the U.S. banking and financial systems were sound
and that the Federal Reserve would act as a lender of last resort in case of a major financial crisis in the United States. The reliance on this system was so widespread that countries with less sophisticated banking and financial systems were often advised to follow similar paths of deregulation and openness to innovation in financial services. In fact, it was precisely the excessive reliance on this system that contributed to its extreme leverage and ultimate collapse.

Most likely, the U.S. financial system, albeit on the mend, will not regain the same international role and relevance that it had during the credit boom of the last two decades. The spillover effect on the real economy and dramatic decline in world trade have raised the specter of deep recession followed by mild recovery. The severity of the recession is already evident in the collapse of world trade and slowdown in international capital flows. The Great Recession has changed the debate from how to manage globalization to how to prevent deglobalization. It has also sparked debates about the future of market economies and capitalism in general and about whether government presence will increase and rein in the instability of markets.1

One of the consequences of this state of affairs is a renewed interest among scholars and policy makers in the study of past financial crises, their origin, unfolding, and consequences. The vast literature on the subject ranges from historical overviews to analyses of the determinants of specific types of financial crises (generally divided into banking, currency, debt, and combinations of the three) to case studies (by country, region, sectors, and financial instruments). The most influential contributions include a combination of these approaches, as, for instance, in the now classic Kindleberger’s *Manias, Panics, and Crashes* (Kindleberger 1989; Kindleberger and Aliber 2005) and more recent Reinhart and Rogoff’s *This Time Is Different* (Reinhart and Rogoff 2009).2

In this paper I focus on two historical cases that show why, at a time of major financial and economic crises, it is important to focus on solving problems in the underlying structure of the economy. Crises often emerge because countries fail to adapt their economies to changes in the structure of the international economy, including the inability of their institutions to cope with such changes. That was the case in the Kingdom of Naples in the seventeenth century, as I explain in the first section. I focus in particular on the work of a contemporary observer, Antonio Serra, against the backdrop of a series of financial and economic crises that plagued the Kingdom of Naples for several decades. In contrast to the mainstream interpretation of the time, which emphasized the monetary origins of these crises, Serra identified “real” factors—from an absence of manufacturing and “entrepreneurial spirit” to a lack of credit and good governance—as the primary causes of the kingdom’s series of crises. With this example, I show how, in a distant past, debates about economic crises revolved around their causes as well as their solutions, with a clear distinction between “monetary” (including exchange rate) and “real” factors. The policy implications of this dichotomy were evident then as now, contributing to the emergence of analyses that, by focusing on a country’s economic structure and position in the international economy, seek to promote long-term growth and reduce the exposure to crises.
The focus on real rather than simply financial factors at times of crisis is also central to another seventeenth-century analysis, which I present in the second section of this paper. In this case an Italian mathematician, Geminiano Montanari, challenged the common practice of blaming financiers and financial innovations for all evils at times of financial crisis. According to Montanari, it was the distortion and misallocation of resources and human capital caused by excess profits in financial activities that contributed to the decline in investments in manufacturing and commerce in the economies of the Italian city-states and ultimately to their demise. He concluded that as the contemporary virtuous example of Holland showed, a country could achieve more sustainable growth by using finance to support investments in real activities, while being less exposed to the vagaries of financiers’ practices.

In the third section, I briefly outline one of the lessons from the Great Depression, which has important implications for the current debate on the design of policies to weather the Great Recession and of exit strategies. On the basis of existing literature, I stress how lack of coordination at the international level contributed to a chaotic demise of the gold standard, the implementation of “beggar-thy-neighbor” policies, and an uneven recovery across the world. Countries that abandoned the gold standard first—and thus were able to devalue and implement expansionary policies—recovered more rapidly than those that broke their “golden fetters” later in the decade (Eichengreen 2008). Within the former group, countries such as Brazil, Great Britain, Colombia, Germany, and Japan, which implemented import substitution policies and channeled investments in real activities, recorded impressive rates of growth (Feinstein, Temin, and Toniolo 2008).

The economic history of the 1930s also shows that lack of international cooperation and the widespread adoption of “beggar-thy-neighbor” policies bore enormous geopolitical and economic costs later in the decade and in the early 1940s. The lesson from this example, in line with the two previous, is that historically, major crises have provided opportunities to rethink a country’s economic structure. Given the current integration of the world economy and the current account imbalances that preceded the Great Recession, perhaps one should consider the world as a large economy and rethink the role of countries as that of regions within one nation. We need to “think outside the box” and adjust long-standing practices, institutions, policies, and international agreements to reflect the new structure of the global economy.

As a historian dealing with the daunting title of this plenary session, “The Road Ahead to a Sustainable Global Economic System,” I chose a few examples from a distant past that, despite differences in the underlying economic structure, historical context, and origin of crisis, give a sense of how certain features of crises are, as Charles Kindleberger has shrewdly remarked, “hardy perennial.” History can provide examples of what did or did not work in specific historical settings, but, more important, it can also teach us to seize the moment, to sense the relevance of the times we are living through. It is during uncertain economic times like the present that governments are required to act decisively and to implement reforms that could set the world economy on a more sustainable path of growth.
The renewed interest among scholars and financial institutions such as the International Monetary Fund in the history of money, financial institutions, and crises, as witnessed by the recent publication of several studies and the formation of new data sets, is a positive development. As well put by Larry Neal in a recent paper, there is a tendency among decision makers in the past as well as today “to spurn the insights developed by historians, feeling that they are irrelevant in the context of modern technology and institutions” (Neal 2009), or, as Gregory Mankiw remarks in assessing the impact of the current crisis on the future teaching of economics, the study of financial institutions “will need to become more prominent in the classroom” (Mankiw 2009). The hope is that decision makers and the financial sector will learn from past mistakes and pay attention to the lessons from history.

Money Is Not the Problem

At the beginning of the seventeenth century, the Kingdom of Naples—a viceroyalty under Spanish control—experienced a series of economic crises whose origins sparked lively debates among Neapolitan and Spanish administrators, merchants, foreign merchant-bankers, and observers of political and economic events.

Over the course of the previous century, the population of the kingdom had grown so much that Naples, with its population of 300,000 at the start of the seventeenth century, rivaled Paris and London, the largest cities in Europe. Population explosion meant an increase in the consumption of commodities such as wheat, wine, and oil that, together with silk, had traditionally constituted the largest share of Neapolitan exports. Increased consumption combined with a decline in manufacturing activities during the same period translated into a worsening of the trade balance and a net outflow of silver and gold. The primary role that agriculture played in the kingdom also meant that bad harvests, such as those occurring in 1593, 1595, 1598, and 1607, weighed heavily on Neapolitan finances because they meant a reduction in export revenues and an increase in payments for the import of wheat and other basic commodities.

The existence of a large public debt exacerbated the precarious financial situation of the kingdom, in particular since most of the interest on the debt was paid to foreign merchants and bankers, thus increasing the outflow of gold and silver coins. The Neapolitan government tried to cope with the ballooning public deficit by rescheduling and negotiating lower interest on the existing debt, but with mixed success. Other attempts to reduce the outflow of precious metals for the repayment of the debt and to inject liquidity in a kingdom plagued by money shortages also failed. Thus the shortage of money and rapid depreciation of the exchange rate were major symptoms of a deteriorating economic situation.

In the early modern period, the exchange rate was at the center of a complex system of international payments involving the conversion of multiple metal, gold, and silver currencies as well as money of accounts throughout Europe. The exchange rate not only enabled the conversion of foreign prices into domestic ones and vice versa but also contributed to the widespread use of credit instruments such
as bills of exchange, which by the end of the sixteenth century had become the most widely used instrument of international payments in Europe. Bills of exchange incorporated the concept of the exchange rate to enable the settlement of accounts among merchants from different states. The existence and diffusion of bills of exchange constituted an expansion of credit (without the physical transfer of metallic money) and contributed to the development of trade and the transfer of wealth throughout Europe (De Rosa 1994; Rosselli 2000). However, over time they became speculative instruments that enabled bankers to amass large profits. This transformation was facilitated by the fact that a handful of international bankers, mainly from Genoa, exercised a de facto monopoly on the international market of these products. These bankers not only held a network of banks throughout Europe but also were in control of the organization of international fairs where a restricted group of European bankers met every three months to settle their accounts, including bills of exchange.

Decrees and treatises from the first decades of the seventeenth century show how monetary issues worried the Neapolitan government as well as experts of the time. Gio Donato Turbolo, an officer of the mint, computed that of the 13 million ducats minted in Naples during the period 1599–1629, only 3 million were still in circulation in the late 1620s. More worrisome for the observers was the poor quality of the circulating money. Attempts to address this problem through a complete overhaul of the monetary system, such as the reforms of 1609 and 1622, failed. Other attempts to inject liquidity into the state, such as a series of debasements (depreciations), resulted in a devaluation of the Neapolitan coinage by 30 percent in the course of the 1610s (Calabria 1991).

Dismissing objections that an appreciation of Neapolitan currency would hurt trade, advisers to the Neapolitan government believed that an aggressive revaluation of the Neapolitan currency would solve the monetary and economic problems of the kingdom. For one of these advisers, the businessman De Santis, the kingdom’s exports were so vital to the life and well-being of people in other states that higher prices would not deter foreign demand. Similarly, the appreciation of the Neapolitan currency would attract foreign capital because foreigners would perceive the Neapolitan securities as less risky and more trustworthy (De Rosa 1994). These views held a strong influence on the Spanish administrators of the Kingdom of Naples and were incorporated into a series of economic reforms that not only failed to tackle the liquidity and credit crisis but also contributed to the worsening of the trade balance.

These views and policies came under attack, in particular in a short treatise, Brief Treatise on the Causes Which Can Make Gold and Silver Plentiful in Kingdoms Where There Are No Mines, written in 1613 by Antonio Serra, a doctor confined to the prison of Vicaria in Naples under indictment of counterfeiting (Serra [1613] 1994). Serra challenged the dominant view that the kingdom’s economic crisis was the consequence of monetary disorders and that monetary and fiscal measures, including manipulation of the exchange rate, could suffice to address the crisis. In his view, the level of the exchange rate was the consequence rather than the cause of monetary shortages (Monroe 1924; Grilli 2006).
To support his views, Serra presented an ideal model of an economy and compared it to the reality of the Kingdom of Naples. For Serra the economic success of a country was dependent on growth in manufacturing and agriculture, an internal process in which a series of factors, including quality of people and good government, played a crucial role. Serra supported his claim by pointing to the lasting economic success of the Venetian Republic. Venice had the most advanced manufacturing sector of the time, the result of investments and ability to retain “the most skilled workers of Europe,” excellent trading capabilities including a credit and monetary system that was the envy of the rest of Europe, and an efficient administration. According to Serra, the combination of these factors had ignited a self-reinforcing process of growth, a virtuous cycle with manufacturing at its center.

The good esteem of the Venetian economy and government was reflected, according to Serra, in the low interest on public debt (4 percent a year), which was in contrast to interest of 8–10 percent in Naples. To make matters worse, foreigners held a large portion of the Neapolitan public debt, which represented a constant drainage of reserves (via service on the debt) and, given the deficit in the balance of trade, a constant increase in consolidated debt. In contrast, Venice’s current account surpluses contributed to the public treasury.

The success of Venice could be ascribed to several factors. The Venetian government had removed major “impediments” to the smooth work of trading and manufacturing activities and created opportunities for merchants and artisans to develop their own businesses—by providing the right incentives and a good environment in modern economic terms. But for Serra, institutional continuity was the most powerful explanation of Venice’s success. Monarchs change, and with them objectives and policies, while in the Venetian Republic the commonwealth had been consistently pursued over time through a constant improvement in the working of various institutions, such as the Senate, and in the selection and management of magistrates, administrators, and officials. The interaction among institutions and across generations (as, for example, in the Senate, where old and young senators learned how to cooperate) and the development of institutional mechanisms (as, for instance, the passing of laws in the Senate, which required a majority of votes) guaranteed the necessary institutional stability and yet allowed for the frequent and necessary renewal of the governing bodies.

Venice represented the model for what to do in Naples, where structural economic and political problems had contributed to the economic decline and monetary crises. For Serra, all the measures suggested by Neapolitan officials—an overvalued exchange rate, a ban on the export of money, and increased taxes on foreigners—went in the wrong direction because they were based on a faulty assessment of the economic reality of the kingdom. He supported government intervention in the real sector of the economy, in particular in manufacturing, but minimal intervention on the monetary side: no exchange rate manipulations, no barriers to capital movements.4

The Brief Treatise ends with a list of proposals to help sovereigns and policy makers in their “difficult task” of steering the economy of their states. Rather than seizing the income realized by foreigners in the kingdom or appreciating the exchange
rate, Serra supported the overtaking of foreign-controlled production by local entrepreneurs, in a slow manner to avoid any disruption in trade. But to achieve this goal—through what today would be called an import substitution strategy—it was necessary to invest, “to introduce in the kingdom” those factors that would support the development of manufacturing and devise regulations that would attract foreign capital and skilled labor. Serra believed there was room to change the course of events, the economic position of a country, with good government and good policies. It was this capacity to change the fate of a country that made “good government” the most powerful of all factors and yet “the most difficult and unpredictable of all” (Monroe 1924).

This case study shows how, at times of major economic crisis, it is important to think beyond the containment and to question the economic structure of the economy in which the crisis is taking place. Contrary to mainstream interpretations of the time, Serra argued that only a complete overhaul of the economic system would prevent the recurrence of crises in Naples. This lesson is still meaningful today. The famous Japanese “lost decade” can be attributed in part precisely to the unwillingness, at least in the 1990s, to undertake radical reforms in the banking sector as well as to boost domestic consumption. The ongoing crisis in the United States can also be attributed in part to the postponement of reforms in the U.S. economy, despite the warning signals at the beginning of the decade, including the dot-com bubble, the Enron collapse, and the rise of unsustainable twin deficits.

At times of major financial crisis with fallout on the real economy, governments should pay attention to analyses of the crises that challenge the structure of the underlying economy, the theories behind them, and the existence of interest groups whose main objective is to maintain the status quo (Johnson 2009). It was in such an environment that Serra advanced his interpretation of the Kingdom of Naples’s economic crisis. His views fell on deaf ears, and the kingdom continued to suffer from recurrent crises. After presenting his views before the viceroy and his cabinet and failing to convince them, Serra was sent back to prison where he died in 1617. In contrast, when monetary and financial troubles emerged in Venice and Holland later in the course of the same century, these governments managed to contain them and to prosper using strategies that would have been very familiar to Serra.5

**Blame the Bankers**

Throughout time, public esteem of money traders, bankers, financiers, and financial institutions has fluctuated between admiration and loathing, the latter sentiment prevailing at times of crisis. Today, public opinion in advanced and emerging economies blames Wall Street, banks, hedge funds, and credit-rating agencies for having issued, backed, and flogged around the world arcane financial products such as mortgage-backed securities. Those who were considered “masters of the universe” are now shunned and ridiculed; financial products that were heralded as “innovations” are now called “toxic assets.” Likewise, in the eighteenth century speculative bubbles in various asset categories (such as extravagant schemes involving future prices of tulip
bulbs in Holland, state-backed securities in France, and trading of company shares in France and England), and the consequent financial havoc gave rise to public scorn of bankers, institutions, and administrators involved in these debacles. Contemporary satirical representations went from commiseration for the victims of these debacles, such as in the plate “Memorial Arch Erected at the Burial Place of Ruined Shareholders” (figure 1), which depicts the rise and fall of John Law’s Mississippi scheme in France, and in “Many Became Crazy Because They Believed in Schemes” (figure 2), an allegory of speculators’ fate. Other plates describe the peril of trading securities with no real backing other than wind (figure 3) or the public scorn of failed stock jobbers (figure 4).6

FIGURE 1.
Memorial Arch Erected at the Burial Place of Ruined Shareholders

The sentiment against financial speculation and its negative impact on the real economy were aptly described by an Italian mathematician, Geminiano Montanari (1633–87), in his analysis of the relationship between financial crises and economic decadence of numerous Italian states in the seventeenth century. Montanari was professor of “mathematics and astronomy” at the University of Bologna and later professor of “astronomy and meteors” at the University of Padua. Like other seventeenth-century scientists—Isaac Newton, for instance, was master of the Royal Mint—Montanari became increasingly interested in money, credit, and finance. In addition to writing a series of treatises on money, he advised the government of the Republic of Venice on the reorganization and management of the mint.

In his comments on the causes of the Italian economic decline in the seventeenth century, Montanari noticed how bankers and foreign exchange traders had enjoyed a high “rating approval,” at a time of economic prosperity in the previous century, despite the church’s moral condemnation of profit. In contrast, financial activities became ostracized when Italian merchant-bankers lost their leadership in commerce and credit to Northern European states. Likewise, he remarked how arbitrage activity had become
FIGURE 3. The Wind Buyers Paid in Wind, or Those Who Are Last Will Remain Hanging on

Source: Het Groote Tafereel Der Dwaasheid [The Great Mirror of Folly] (Amsterdam 1720), from Rare Book, Manuscript, and Special Collections, Duke University.
so pervasive that “all throughout Italy the wealthiest merchants” were getting wealthier by investing most of their resources and time in financial speculations (Montanari [1683] 1804).

But for Montanari, it was unfair to consider bankers as the main cause of recurrent monetary crises and general economic decline. In fact, they had switched their capital from manufacturing to finance in response to increasing costs of labor and uncertainty about the future of Italian cities’ economies. Financial activities enjoyed higher returns and fewer risks than “manufacturing of textiles” and commerce of “silk, spices, wool, and other commodities,” sectors with a “much more complicated organization” and in which “most of the profit would go to workers.” Businessmen acted rationally since activities like arbitrage enabled them to maximize profit and minimize risk. They achieved this goal by exploiting information advantage, such as knowledge of credit instruments, and implementing shrewd schemes. Montanari ([1683] 1804, 151) described this state of affairs, a picture that bears an interesting resemblance to more recent cases of financial speculation, in the following passage:
I personally praise the shrewdness of those who having only to consider their own personal return choose the category of business that produces the quickest and least risky profit; and I say that they commit no crime by standing with their eyes wide open ready to spot new opportunities in any place [and] to profit from sending coins in exchange for other coins; nor by maintaining correspondence and keeping informed about any public deliberation and decree related to monetary affairs enabling them to promptly exploit with their sharp arithmetical approach any possible exchange of different types of coins; I even admire when in order to conduct their transactions they diligently and rapidly borrow money at interest from others, or when they stay particularly vigilant about neighboring markets.

These speculators followed a particular sequence to lure investors who wanted to make ready profits. First, traders identified and gathered in neighboring states coins that they knew commanded a greater premium in their own market. They introduced these coins in their own market and supported their “credit” at a value higher than the official price by using and accepting them even when that entailed a loss. With these actions, traders persuaded city dwellers and artisans to accept these coins. At this stage, speculators “readily import[ed] a large quantity of these overvalued foreign coins from the state where they [were] minted,” and before authorities could intervene, in general with a decree banning the circulation of such coins, they had flooded the entire state while hoarding and exporting the best coins—a typical case of Gresham’s Law.

Returns on monetary financial activities were so high that governments could do very little to prevent the constant drainage of capital and talents. Money traders and bankers could easily double the initial capital in one year, “not a small gain” in Montanari’s words, and could even increase their profits by “borrowing from others.” This behavior, though justifiable from speculators’ point of view, had devastating consequences for Italian cities, the centers of earlier Italian success, since manufacturing and commercial activities “used to employ half the population,” while trading in money and credit instruments “employed few people” (Zanalda 2009).

Italian merchants’ behavior had to be seen in the larger context of Italian decline during the seventeenth century. The gradual rise of nation-states such as France and the growth of trading powers such as England and Holland, Italy’s sporadic participation in the new trading routes in the Atlantic and in Asia, internal divisions, wars, foreign occupations, and territorial struggles still plaguing the peninsula during the seventeenth century contributed to the decline. Montanari also identified other specific factors: lack of government support, migration of skilled artisans, struggles and disputes between merchants and workers, guild power and regulations, fiscal pressure (duties) that hampered commerce, and the tendency among wealthy merchant-bankers to invest their capital in land and estates rather than in productive activities. On the latter, Montanari remarked that by “investing their capital in earldoms and marquisates . . . conducting the leisurely life of princes” while breeding “distaste for the exercise of the once esteemed merchant profession,” this “urban mercantile nobility” transformed the common perception of the merit and social standing of commercial activities, inflicting a “fatal blow to manufacturing and commerce in most cities” (Zanalda 2009).
In a few passages Montanari summarized the main structural problems that plagued Italian economies over the course of the century. He had lived through what historians now consider the most critical period, 1620–80, for the Italian economy of the early modern era. The lesson was clear: several factors, including excessive speculation on credit and financial instruments, had contributed to the misallocation of resources and the loss of competitiveness of Italian economies. As has been the case in several instances throughout history, it was a self-inflicted wound. Credit and financial instruments that had been invented to expand credit, reduce transaction costs, and spread risk—with a positive impact on real activities—had become increasingly the object of speculation, detached from their original purpose.

Then as now, a return to a more responsible use of financial innovation is a condition sine qua non to enhance the availability of credit and efficient allocation of capital. Calls for new regulations and a more conservative use of financial leverage both at the domestic and international levels are part of an ongoing discussion in the United States, in the European Union, and at the G-20 level.

The other related challenge for policy makers is to restore trust in the financial system, while not hindering innovation in finance, an essential feature of market economies. A generalized backlash against finance and sometimes, by affiliation, against capitalism—whether in eighteenth-century Europe, during the Great Depression, or now—could stifle the introduction of new financial and credit instruments. For instance, resistance to the introduction of new credit instruments after the collapse of John Law’s system in the early eighteenth century might explain why France’s transformation into a commercial society was slower than that of England and Holland (Atack and Neal 2009). Likewise, the association between financial speculation and capitalism has often characterized public outcry against bankers and finance at times of crisis, whether in Germany in the 1920s, in the United States and Europe in the 1980s, or in Asia in the 1990s when, “in the midst of its homegrown financial crisis, capitalism as practiced in that continent was everybody’s favorite punchbag” (Pilling and Atkins 2009).

Off Gold

The current crisis has often been compared to the Great Depression, and policy makers, in particular in the United States, have been trying to avoid the policy missteps that exacerbated the downturn after the stock market crash in October 1929. The chairman of the Fed, Ben Bernanke, one of the foremost experts on the Great Depression, has often referred to events and policies of that period in testimonies and speeches and acted to avert a global financial meltdown. Among the numerous lessons from the 1930s that can be drawn from a vast body of research, I focus here on a brief overview of the different paths to economic recovery experienced by countries after the abandonment of the gold standard. The relevance of this decision for the recovery process has been analyzed with different nuances in various studies (Kindleberger 1986; Temin 1989; Bernanke 2000; Eichengreen 2008; Feinstein, Temin, and Toniolo 2008).
After the 1929 crisis, policy coordination among the United States and European nations would have enabled the implementation of a coordinated program of macroeconomic reflation, lower interest rates, and expanded money supply in all countries, with the result of stimulating economies without destabilizing exchange rates. Lack of cooperation among governments and their central banks instead characterized countries’ response in the early 1930s, which in turn generated further deflationary pressure on the world economy and exposed weak currencies, mainly the pound sterling and the mark, to speculative attack (Feinstein, Temin, and Toniolo 2008). Central banks and governments believed in the monetary orthodoxy of the gold standard, which prevented countries from embarking on countercyclical policies and in some cases, such as in the United States and France, intensified the economic downturn (Temin 1989). Given the lack of coordination and urgent need to implement expansionary monetary and fiscal policies, some countries, like the United Kingdom in 1931, took the step of going off gold and embarked on a program of unilateral reflation. This implied that countries improved their economies at the expense of other countries, through what it is known as “beggar-thy-neighbor” devaluation. Once off gold, the Bank of England lowered interest rates and devalued the currency, with beneficial effects on the trade balance. Scandinavian countries took the same step in 1931 as did countries with colonial or trade (like Argentina) relationships with England, which went off gold and pegged their currencies to the pound sterling. Germany also abandoned the gold standard, but only after Hitler rose to power and the government adopted expansionary policies including a vast program of military expenditure, while maintaining controls on capital movements. Latin American countries abandoned the gold standard the same year. The ensuing expansionary fiscal and import substitution policies helped most countries in the region to recover rapidly (Feinstein, Temin, and Toniolo 2008).

Only after experiencing the devastating effect of the financial, banking, and economic crises in 1931–32, the United States abandoned the gold standard in 1933. The devaluation of the dollar, fiscal and monetary expansionary policies, restructuring of the banking sector, together with other New Deal measures helped the recovery of the U.S. economy until 1936 (Kindleberger 1986; Bernanke 2000). That year, the Fed, concerned about future inflation, began to withdraw liquidity, while President Roosevelt, concerned with the ballooning federal budget deficit, supported a tightening of fiscal policy through tax increases and spending cuts. The combination of tight monetary and fiscal policies transformed the fiscal deficit of 1936 into a surplus the following year but also pushed the United States back into recession—the real GDP contracted more than 3 percent in the next two years. That experience explains why in the United States today numerous economists inside and outside the administration are advising the Fed and President Obama to dismiss calls for a tightening of monetary and fiscal policies because of the apparent risk of inflation down the road (Blinder 2009).

France and other countries of the “Gold Bloc” (Italy, Belgium, Switzerland, Poland, and the Netherlands) maintained the gold standard until 1936. Until that year, these countries followed deflationary policies that hampered the recovery process. Even worse, trade among members of this bloc was hampered by the over-
valuation of their gold standard parities. After going off gold in 1936, these countries devalued their currencies and began to recover.

Overall, the best solution for the world economy would have been a coordinated effort among countries to dismantle the gold standard in an ordinately manner. This did not happen, and countries that began to use exchange rate devaluation early in the decade together with other strategies such as expansionary policies (United Kingdom, Sweden, and Japan), protection and import substitution (Brazil and Colombia), and capital controls and domestic expansion (Germany) outperformed those that abandoned the gold standard a few years later (the United States, France, and the Gold Bloc countries). Table 1 summarizes with “a good dose of oversimplification,” the economic recovery paths that followed the abandonment of the gold standard and the implementation of other policies (Feinstein, Temin, and Toniolo 2008).

### TABLE 1. Exchange Rate Policies and Paths to Economic Recovery in the 1930s

<table>
<thead>
<tr>
<th>GDP per person: 1929 = 100</th>
<th>1929</th>
<th>1932</th>
<th>1935</th>
<th>1938</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Policy and Country</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early devaluation and</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>domestic expansion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>100.0</td>
<td>93.5</td>
<td>105.0</td>
<td>113.9</td>
</tr>
<tr>
<td>Sweden</td>
<td>100.0</td>
<td>94.8</td>
<td>109.4</td>
<td>122.1</td>
</tr>
<tr>
<td>Japan</td>
<td>100.0</td>
<td>96.8</td>
<td>104.6</td>
<td>120.8</td>
</tr>
<tr>
<td>Early devaluation, protection, and import substitution</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>100.0</td>
<td>89.5</td>
<td>101.1</td>
<td>112.2</td>
</tr>
<tr>
<td>Colombia</td>
<td>100.0</td>
<td>100.4</td>
<td>111.4</td>
<td>122.5</td>
</tr>
<tr>
<td>Controls on capital movements and domestic expansion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>100.0</td>
<td>83.0</td>
<td>101.7</td>
<td>123.3</td>
</tr>
<tr>
<td>Italy</td>
<td>100.0</td>
<td>95.3</td>
<td>101.8</td>
<td>107.2</td>
</tr>
<tr>
<td>Central planning and autarky</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soviet Union</td>
<td>100.0</td>
<td>103.8</td>
<td>136.3</td>
<td>155.1</td>
</tr>
<tr>
<td>Late devaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>100.0</td>
<td>71.1</td>
<td>77.5</td>
<td>87.0</td>
</tr>
<tr>
<td>Gold Bloc (deflation and late devaluation)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>100.0</td>
<td>84.0</td>
<td>86.8</td>
<td>94.8</td>
</tr>
<tr>
<td>Belgium</td>
<td>100.0</td>
<td>91.1</td>
<td>96.8</td>
<td>95.6</td>
</tr>
<tr>
<td>Switzerland</td>
<td>100.0</td>
<td>90.2</td>
<td>93.3</td>
<td>100.9</td>
</tr>
</tbody>
</table>

The lesson is that at times of major financial crises, including currency crises, governments should keep an open mind with regard to solutions that go against the orthodoxy of the time. Postponing decisions, such as abandoning the gold standard in the early 1930s or maintaining unrealistic exchange rates, as happened during the financial and currency crises of the 1990s, worsens the crisis and hampers the recovery process. In the latter case, currency devaluation and the establishment of flexible exchange rate regimes contributed to contain the crisis and reassert within a few years the strength of Asia’s economies (Eichengreen 2008).

Conclusions

In the end what can we say about the road ahead after the current crisis? How can we prevent a world “lost decade” or a Great Depression redux and put the global economy onto a sustainable path of growth? Peter Temin, in a study published before the current crisis, argued that the 1990s had features of a postwar decade such as the 1920s and 1950s. He worried that given the size and type of pre-crisis problems—stock market exuberance, excessive financial leverage, and international imbalances created by international differences in saving and spending behavior—the 1990s looked very similar to the decade that preceded the 1930s (Temin 2006). Events have proved him right, at least for the initial relevance and global reach of the crisis. It is hoped that because of our knowledge of the 1930s and other crises—overall the depression became “great” to a large extent because of the initial (and again in 1936) mismanagement of the crisis in the United States, the absence of stabilizers, and the lack of international cooperation—we will not look back at the 2010s as a new 1930s. On the positive side, governments around the world have been collaborating to prevent a complete meltdown of the global financial system. G-20 meetings, increased reliance on (and funds to) the International Monetary Fund, implementation of large fiscal stimuli, and expansionary monetary policies have, for the time being, signaled a willingness to stabilize the world economy and prevent a rush to “beggar-thy-neighbor” types of policies. However, this encouraging trend can be readily reverted if the steep downturn in both world trade and capital flows of the last year continues, and the United States, Japan, and other Asian and European countries are not able to maintain expansionary policies.

To achieve sustainable growth at the global level, it is important to apply one of the main lessons stressed throughout this paper: governments should consider this crisis as an opportunity to implement reforms, even structural reforms, and, as Martin Wolf puts it, “adapt the market economy to their own traditions” (Wolf 2009). This does not imply a rejection of the great achievements of the last two decades, but an effort to recapture and sustain the rapid growth of the world economy without recreating the same type of global imbalances.7

Likewise it is important to address the problems that have emerged in financial sectors in the United States and Europe. As in several historical examples, the level of leverage and profits and the lack of supervision and transparency of the banking and financial systems (this time around in the United States and Europe) are used to indict
the excesses of the prevailing form of capitalism. Influencing the debate on the future of Western economies are the U.S. origin of the current crisis together with the realization that Asian economies, excluding Japan, have succeeded in recovering from past crises on the basis of a softer version of capitalism characterized, among other things, by more protected credit systems and a greater focus on the real economy. In essence, the pendulum is swinging back toward a market economy in which the government plays a greater role and maintains a stricter control on private finance. This swing has different nuances—in particular, with regard to use of the term “socialism”—similar to what Peter Temin argues in Lessons from the Great Depression, “If there is a renewed depression . . . then we should expect a swing of the policy pendulum back toward socialism. Capitalism thrives during economic stability. It wilts in depression. Socialism appears to be the reverse. It fades during stability . . . but it flowers in depression with its support of economic planning and distribution of social dividend” (Temin 1989, 136).

Where do we go from here? In line with the idea that it is time to adopt bold actions to prevent a globalization backlash in both trade and capital and the recurrence of new crises, it is essential to address the imbalances between China and the United States. Signals in both countries are encouraging. The U.S. president and his administration seem committed to implementing reforms that will have a profound effect on the U.S. economy, although all attempts could derail if the U.S. fiscal position becomes untenable. In contrast, China, to make up for declining global demand, has begun to implement measures and commit resources to create inner dynamism as a complement to export-driven growth (Kynge 2009).

The fate of the U.S. dollar represents an important issue for the stability of the world economy. Again, this area is open to bold proposals for reform, such as the one recently advanced by the governor of China’s central bank, Zhou Xiaochuan. Zhou questions the long-run sustainability of an international monetary and financial system with the U.S. dollar at its center. An alternative, according to Zhou, would be to revise an old Keynes idea of an international reserve currency, another version of special drawing rights. This is an interesting proposal, an attempt to answer a legitimate question about the future role (and value) of the U.S. dollar as well as that of the China renminbi. This ought to be addressed sooner rather than later to avoid the devastating effects on the global economy of either a collapse of the dollar or a dramatic increase in interest rates, which would hamper the recovery process in both advanced and emerging economies.

In their recent analysis of eight centuries of financial crises, Reinhart and Rogoff conclude that severe financial crises have deep and lasting effects on asset prices, employment, and output. On average, housing prices decline for six years, unemployment rises for five, and output declines for two. Massive increases in government debt are the norm at the end of recessions created by financial crises (Reinhart and Rogoff 2009). It seems that the current crisis is following the same path, at least in the United States, Europe, and Japan. Asian economies as well as Brazil seem, at the moment, poised to emerge in a stronger position for reasons that bring us back to issues previously discussed. In a recent interview, Kishore Mahbubani, dean of Singapore’s Lee Kuan Yew School of Public Policy, argues that the current form of Asian capitalism is
the result of Asians having adopted the basic features of Western capitalism, such as reliance on free markets, navigated through the 1990s crisis, listened to International Monetary Fund advice, and added their own lessons, which include “do not liberalize the financial sector too quickly, borrow in moderation, save in earnest, take care of the real economy, invest in productivity, focus on education.” To which he added, “While America was busy creating a financial house of cards, Asians focused on their real economies” (Pilling and Atkins 2009). It is a sign of the times that the economic agenda President Obama is trying to implement in the United States focuses on similar principles or maybe it is simply, as I show in the case of Naples in the seventeenth century, that at times of crisis governments are forced to address the foundation of their economies (Leonhardt 2009). And as Serra pointed out in the case of Naples in the seventeenth century, the “quality” of the government is a crucial factor, including its ability to implement and sustain reforms. Hence, the stability of the global economic system will depend in large part on the ability of the United States and Europe to reform and the ability of emerging economies to manage a new global economy in which they will have a greater economic and political role.

Notes

1. The origin and dramatic unfolding of the Great Recession in advanced economies has also raised questions among economists and policy makers about the suitability of current economic theories and models that have been constructed for a different world. As well put in a Financial Times editorial, “Most models depict economies kept close to equilibrium by smooth adjustments. But we face a very real danger of large, abrupt changes, bank collapses, or currency crises. And unlike what most models assume, prices are not properly clearing all markets” (Financial Times 2009).

2. For further examples of historical overviews, see, among others, Garber 2000; Bordo and Eichengreen 2002; Eichengreen 2002; Neal and Weidenmier 2003; Caprio, Hanson, and Litan 2005; Ferguson 2008.

3. Although Holland experienced one of the first recorded cases of asset bubbles in the second half of the seventeenth century, the famous Tulipmania, the resilience and depth of the Dutch financial system of the time contributed to lessen the impact on the real economy (Garber 2000).

4. Serra opposed all forms of monetary alterations, in particular debasement, but also conceded that contrary to general opinion, the circulation of coins with no intrinsic value did not affect the volume of commerce within a state. This observation carried an important lesson about the relationship between quantity of money and economic activity. Carlo Cipolla, and more recently Thomas Sargent and François Velde in their research on the role of the parallel circulation of coins with intrinsic value (silver and gold money) and low or no intrinsic value (alloy coins and later paper money), concludes that the vast circulation of devalued coinage provided the necessary liquidity for the rise of market economies in European states after the fourteenth century (Cipolla 1958; Sargent and Velde 2002).

5. The novelty of Serra’s contribution was singled out by Joseph Schumpeter. The Austrian economist praised Serra for his demonstration that natural resources, quality of people, industry and trade, and the efficiency of government, more than money, determined the success of production and commerce and that “if the economic process as a whole functions properly, the monetary element will take care of itself without requiring any specific
therapy.” In Schumpeter’s view, for decades to come, “There [had been] nothing like this [analysis] anywhere” (Schumpeter 1954).

6. All these plates come from The Great Mirror of Folly, a famous early eighteenth-century volume published in Amsterdam (Cole 1949).

7. This can be seen as a continuation of the debate about the benefits and costs of the globalization process. It can be reduced to two main positions, as summarized by Andrei Shleifer. On the one hand, there are those, like Stanley Fischer, who believe that market forces, open economy, macroeconomic stability, and good institutions explain rapid economic growth in emerging economies—a positive view of the process. On the other hand, there are those, like Joseph Stiglitz, who criticize free market policies and advocate a greater role for the state, extensive regulations, and some form of capital control—a view that globalization has to be managed and customized to countries’ traditions (Shleifer 2009). Perhaps the crisis will help to reconcile these two views.

References


Comment on “A Sustainable Global Economic System after the ‘Great Recession’? Some Lessons from History,” by Giovanni Zanalda

JEROMIN ZETTELMEYER

Most of the discussion on the future of the global economic system after the Great Recession takes a normative angle. What policy conclusions should we draw from the crisis of 2007–09? How should the crisis affect the structure, size, and internal incentives of financial institutions, for example? What does it imply for regulatory reform and the role of the state in the financial sector? What does it suggest about the role of existing international institutions? Does it give rise to a case for new cross-border institutional arrangements? What does it imply about how to conduct macro-economic policies in large countries, such as the United States and China?

These are important questions to which answers are being suggested and debated, including in several chapters of this volume. But the consequences of the Great Recession can also be debated from another angle that is just as important but has received less attention: What will be the likely outcome of the crisis for the global economic system? Will the outcome lead to a crisis of capitalism or a backlash against the consensus that has been in place at least since the Reagan-Thatcher era of deregulation and privatization? Will it lead to major new reforms that will make the world economic system safer? Or will it ultimately not lead to very much at all: a few reform attempts, perhaps, that either do not go far enough or are based on a misdiagnosis of the problem? This “positive” or predictive angle is important not only because it is important to prepare for what comes next, but also because it could give rise to “meta-normative” ideas—ideas not so much on how precisely the world should change after the crisis, but on how the political economy of reforms should be influenced in order to achieve a good outcome.

What makes Giovanni Zanalda’s paper interesting and timely is that it speaks to this less explored angle. It does so by analyzing three historical crisis episodes, with
clear relevance for today. Zanalda’s message is that in principle, correct diagnoses of these crises were available even while they were still under way. However, the extent to which these diagnoses were put into practice varied widely. In other words, the three episodes seem to suggest that societies have the opportunity to draw the right conclusions from crises, but do not always take them.

**Crises and Reform**

When do societies draw the right lessons from history, and when do they either fail to react or—perhaps just as undesirably—overreact? This question is outside the scope of Zanalda’s paper and perhaps outside the scope of historiography, which is concerned mainly with getting the specifics of what happened right rather than with identifying regularities across many episodes. However, the question is related to a core question of political economy: what circumstances trigger policy reform? See, for example, classic contributions by Krueger (1993) and Rodrik (1996). Reform typically happens incrementally, as the relative strengths of social groups backing or opposing reforms change over time. But reform can also be triggered by large economic or political shocks, which may suddenly change the perceived interests of a particular group or weaken an incumbent. In this sense, crises can represent an opportunity.

From a political economy angle, it is useful to distinguish between three types of societal reactions to crises. One is that nothing much happens, sometimes in spite of an awareness that reform would be desirable. The seventeenth-century fiscal crisis in Naples that inspired Antonio Serra’s treatise, described in Zanalda’s paper, serves as an example. Another reaction is that reforms happen, albeit within the context of the existing political and financial order. There are no examples of this in Zanalda’s paper, but a recent case that comes to mind is a wave of reform (mostly of monetary institutions and fiscal policy) in Latin America, triggered by the crises of the middle and late 1990s. Finally, crises can lead to wholesale reactions against the status quo. The classic example is a revolution set off by a combination of economic hardship and political shocks (the Russian and French revolutions come to mind). Another example is the Great Depression, which is one of the three crises discussed by Zanalda. Zanalda argues that monetary reforms came too late and were too uncoordinated to mitigate the impact of the Depression. This may well be true. Eventually, however, the Depression led to a sea change in the relationship between the state and the market—the birth of the welfare state—and a new international economic order in the form of the Bretton Woods system.

The three types of crisis reactions differ also from a normative perspective. A failure to reform, by definition, is neutral compared to the status quo (although it may be disappointing to those, like Serra, who believe that the status quo is flawed). Successful reforms, again by definition, are an improvement. The third type of reaction—a shift in economic or political systems against the status quo—cannot be categorized ex ante as good or bad. It depends on the nature of the status quo and the system that replaces it.
Reactions to This Crisis: Too Much or Too Little?

Where should we expect the current crisis to take us? Which type of reaction will it trigger?

Zanalda’s implicit fear is that the global crisis of 2007–09 might end like the Naples crisis of 1610: namely, with a failure to take action or at least the right action. This is certainly one possibility. However, based on historical precedent (beyond the three episodes studied in the paper), the opposite also has plausibility; namely, that the ongoing crisis will trigger a reaction, and possibly an overreaction, against the status quo. The crisis has been billed as a crisis of capitalism, coming at the end of an era of relatively unfettered private activity (albeit possibly distorted by public safety nets), particularly in the financial sector. More important, it seems to share some of the traits of previous crises that did lead to large reactions. It has been large, synchronized (in the sense that it has affected many countries at the same time and for similar reasons), and there appears to be a clear culprit: financial capitalism.

From the perspective of developing and emerging-market countries, which are the focus of this conference, three recent precedents come to mind.

First is the Latin American debt crisis of the 1980s. While this crisis was triggered by external shocks, the underlying vulnerability of the region (in the form of high public debt and inactivity of the region to “grow out” of the initial shock) reflected the failure of the state-led development model that the region had adopted in the 1950s and 1960s (Fraga 2004). The crisis led to the replacement of this model by the “Washington Consensus” (Williamson 1990), emphasizing a different and in some ways more limited role for the state (namely, to secure property rights and provide social safety nets rather than to plan and lead the process of industrialization), external liberalization, and a greater role for the private sector.

Second is the Asian crisis of 1997–98, which led to the collapse of incumbent regimes (most famously, the Suharto regime in Indonesia), a backlash against the “crony capitalism” that accompanied these regimes, and a wave of reforms. These included improved financial regulation, better corporate governance, a clearer separation between public and private sectors, and more competition (Kochhar, Loungani, and Stone 1998; Lane and others 1999; Asian Development Bank 1998, 1999, 2000).

Third is the crisis in the Russian Federation and neighboring transition economies in 1998–99. In the political arena, the anti-status quo reaction to this crisis was comparatively muted (for details, see EBRD 2009, ch. 6). However, the crisis had a significant short-term effect on economic reforms—in the opposite direction as the previous two crises. In the five years preceding the crisis, reforms proceeded briskly, with an average of 66 “upgrades” a year on the transition indicator scale of the European Bank for Reconstruction and Development (EBRD) for the transition region as a whole. But in 1998, only 35 transition indicator upgrades were awarded, and 13 downgrades were issued in nine countries (Belarus, Latvia, Romania, Russia, Serbia, Turkmenistan, Ukraine, Uzbekistan, and the former Yugoslavia). With only 34 upgrades and seven downgrades, 1999 was also a bad year for reform (see figure 1).
The three episodes suggest, first, that an anti–status quo reaction during and after the ongoing crisis is plausible. Second, if it does occur, it would most likely be directed against market reform, along the lines of the reactions to the 1998–99 crisis in the Russian Federation. From the point of view of development, this would be highly regrettable, since the reforms share some credit (with a favorable external environment) for generating good growth in developing countries since the late 1990s. Even from a pure stability point of view, the emerging markets have done relatively well in this crisis, in part because of macro-institutions and financial sector reforms since the last wave of crises (EBRD 2009, box 1.2). This is the first major financial crisis that has not hit the developing world harder than the advanced countries, at least since the new era of financial integration began in the early 1970s. Even the transition region, which was exceptionally hard-hit in terms of decline in average output, has proven resilient in the sense that the crisis has not led to currency collapses and financial meltdowns (EBRD 2009, ch. 1).

**Evidence from the Transition Region**

For this reason, the possibility of an anti–status quo reaction to the crisis in the transition region has been a significant concern to the EBRD and other proponents of reform. But what is the evidence? Just over 12 months into the crisis proper—following the collapse of Lehman Brothers in September of 2008 and the sudden stop of capital flows to emerging markets in the last quarter of that year—it is too
early to draw definitive conclusions. However, some initial observations are worth reporting.

Most obviously, the reform reversals that followed the Russian crisis have been largely absent. For the period from mid-2008 until the third quarter of 2009, which coincides almost with the most virulent period of the crisis in emerging Europe, the EBRD Transition Report 2009 identifies only four transition reversals. Moreover, the report emphasizes that at least three of these reversals, which refer to nationalizations in Latvia, Montenegro, and Kazakhstan, can be justified as part of the crisis response strategies of the three countries. Compared to the aftermath of the 1998 Russian crisis, the ongoing crisis in the transition region is noteworthy mainly for what has not happened: a resort to protectionism, capital controls, and nationalizations on a broad scale, let alone sovereign default. While the scope of the state has been expanded in many countries—most notably, in Kazakhstan—the rationale for this has been no different from crisis-related bank nationalizations in the United Kingdom, for example.

But what if the reform reversals still lie ahead of us? If the crisis leads to a significant shift in public opinion—and, via elections, in political power—it could still trigger transition reversals in the future. One source that may provide answers in this respect is the next round of the EBRD–World Bank Life in Transition survey, which will go into the field in the first half of 2010 and allow a comparison with reform attitudes in the region at the time of the first round of the survey in 2004–05. In the meantime, government changes in the transition region can provide some clues. The EBRD Transition Report 2009 (EBRD 2009, box 6.2) examines all elections and political turnover from January 2008 until September 2009 and finds no evidence of a political shift against market reforms. The declared policy stance of governments appears to have either remained the same or (in eight out of the 29 EBRD countries of operations) become more reform friendly. On this basis, it appears unlikely that the region will witness a strong backlash against reforms, at least in the near future.

Conclusions

From the perspective of developing countries and transition economies, the main concern in this crisis is not so much that it may lead to reform inaction, although this is also a concern, particularly with respect to the financial sector, but that it may lead to a reversal of previous market-oriented reforms. This would be a tragedy, in part because many reforms of the last decade have contributed to an improvement in macroeconomic and financial fundamentals that has strengthened the resilience of developing countries during this crisis. Fortunately, evidence from the transition region so far does not support the idea that a backlash is under way.

However, we should not draw too much comfort from this fact. As fiscal needs in the advanced countries absorb a larger share of world saving, private capital flowing to developing countries will likely be more scarce and more expensive in the medium term. This, in turn, will put a premium on new reform in areas such as education, health, and infrastructure that can unlock domestic sources of growth, particularly in
the transition region, which has relied excessively on foreign financing. Generating the requisite fiscal space is likely to require fiscal-structural reforms. With the core supporters of reform—the middle class—under the stress of high unemployment, household debt, and falling housing prices, the appetite for new reform efforts is likely to be weak. Hence, Zanalda’s main worry—that the crisis of 2009 may be a missed opportunity in much the same way as the Naples crisis of the 1610s—could yet come true. Preventing reform from stagnating will require applying the same policy effort to the postcrisis reform agenda as governments and international institutions have applied to the crisis response in the past 18 months.

References


Innovation and Competition
Innovation and Financial Globalization

PHILIP R. LANE

This paper examines the links between international financial integration and the level of innovation activity. If financial globalization boosts innovation, this helps to explain the empirical evidence indicating that increased financial integration conditionally raises the level of productivity and long-run living standards. The paper finds that conditional on the level of development, more integrated economies do exhibit higher levels of innovation activity, but the impact differs across equity-type and debt-type dimensions of international financial integration. Moreover, the gains from equity-type integration kick in for countries at relatively low income levels, whereas the gains from debt-type integration are only found for high-income countries.

This paper asks how international financial integration affects the levels of productivity and innovation activity. This is a timely question, since the current global financial crisis is leading to a reevaluation of the net benefits from international financial integration. Indeed, the data indicate that developing countries have been able to grow without net capital inflows (see, among others, Prasad, Rajan, and Subramanian 2007; Rodrik and Subramanian 2009). Moreover, financial globalization has not generally delivered a more stable path for consumption or output for developing countries (Kose, Prasad, Rogoff, and Wei 2009). However, there is a recent accumulation of evidence that under certain conditions, international financial integration can help to deliver a higher level of productivity, which is the driving force for long-term living standards.¹ Accordingly, this paper examines in more detail one mechanism by which international financial integration may raise productivity, which is its potential impact on the level of innovation activity.

Philip R. Lane is with the International Institute for Integration Studies in the Economics Department at Trinity College Dublin and the Centre for Economic Policy Research in London. He is grateful to Jan Svejnar, the anonymous reviewers, and Richard Tol for helpful comments, to Daniel Lederman for help with data, and to Peter McQuade for excellent research assistance.

Annual World Bank Conference on Development Economics 2010, Global
© 2011 The International Bank for Reconstruction and Development / The World Bank
The potential connection between financial globalization and productivity is critically important. While international financial integration may also operate through capital deepening, the most important potential payoff occurs when financial integration raises total factor productivity. In particular, Gourinchas and Jeanne (2006) show that international financial integration can only offer a limited welfare gain if the level of productivity is unaffected, since the capital deepening effect only serves to bring forward convergence to a country’s conditional steady-state level of output. In contrast, an increase in productivity improves the steady-state level of output, offering a long-term welfare gain.

The paper focuses on innovation activity, since technological progress depends on purposeful efforts to develop new technologies or, especially in developing countries, to move closer to the frontier by adopting existing technologies developed elsewhere. Even in the latter case, the adoption of existing technologies is costly, requiring local research and development (R&D) activity. In addition, the attainment of technological progress typically involves resource reallocation across firms, with higher-productivity firms expanding and laggards being driven out of business. For this reason, economic environments that facilitate such firm-level dynamism may be more conducive to higher rates of effective innovation activity.

The role of investment in knowledge and resource reallocation in driving innovation activity suggests that the financial system has an important part to play in promoting innovation. For instance, the costs incurred in R&D may require external-to-the-firm funding. Similarly, if technological advances are embodied in new firms, this requires a financial system that is able to support the early-stage growth of de novo enterprises. At an industry level, if productivity growth is higher in some sectors than in others, the financial system must have the capability to redirect funding from slower-growing to faster-growing industries. For these reasons, a sizable literature emphasizes the importance of financial development in determining the level and effectiveness of innovation activity.

In turn, the positive contribution of financial development begs the question of whether international financial integration has the potential to boost the level of innovation activity. First, international financial integration may be helpful as a result of the role played by financial globalization in accelerating the development of domestic financial systems in developing countries. Second, the specific characteristics of innovation activity may be especially supported by the entry of foreign investors, in view of the risk profile of the innovation process. Third, one line in the recent literature on financial globalization emphasizes that financial globalization may operate indirectly by tilting the political economy calculus in the direction of improving the general domestic institutional environment (see, among others, Kose, Prasad, Rogoff, and Wei 2009). Through this indirect channel, international financial integration may further boost innovation activity, through the positive impact of institutional reform on the investment climate.

However, the recent literature also finds that the gains from financial globalization are typically conditional on the country’s level of development (Kose, Prasad, Rogoff, and Wei 2009; Bekaert, Harvey, and Lundblad 2009; Masten, Coricelli, and Masten 2008). This pattern has been found in myriad studies of the influence of
financial globalization on output growth, productivity growth, and volatility measures. The precise conditioning variables differ across these studies, but the general theme is common: financial globalization may not be helpful (and can even be harmful) if the domestic economy is not sufficiently developed to marshal the potential gains. Accordingly, it is important to investigate whether such threshold effects are also present in determining the relation between international financial integration and level of innovation activity.

A related point is that the gains from international financial integration may be distributed unevenly across the different types of cross-border investment activity. Most obviously, equity-type investments (foreign direct investment or portfolio equity) are different in nature to debt-type investments (bank loans and deposits or portfolio debt) along several dimensions. Accordingly, the general impact of different types of financial integration on the level of innovation activity may not coincide. Moreover, the threshold level of economic development required to gain from international equity integration may not be the same as that for international debt integration. This paper explores these potential differences.

The paper is structured as follows. It describes the role played by innovation in determining the level of productivity, examines the potential links between financial globalization and the levels of productivity and innovation activity, and conducts an empirical analysis of the cross-country relation between international financial integration and the levels of innovation activity. A final section concludes.

### Innovation and Productivity

There are two sources of productivity growth. First, the global technology frontier is advanced through the development of new or better types of products, plus efficiency gains in the methods of producing the existing range of products. Second, for most countries, the existing level of technology is some distance away from the frontier. For these countries, productivity growth can be achieved through the adoption of superior forms of technology that have already been introduced in leading-edge economies. For developing countries, it is natural that the primary focus has been on understanding how to close the gap between the existing technology and the technology frontier.

In either case, R&D costs need to be incurred in order to develop new technologies or adopt existing technologies. Keller (2004) emphasizes this point: international technology diffusion does not occur in a passive fashion but rather requires purposeful investments in order to acquire and exploit the technological advances made elsewhere. Clearly, this is a multidimensional challenge, and the successful adoption of new technologies is facilitated by factors such as improvements in the stock of human capital, high-quality domestic institutions, and integration into the global trading system.

Moreover, Coe, Helpman, and Hoffmaister (2008) find evidence that these factors interact with each other. For instance, R&D expenditures are more effective the higher is the quality of domestic institutions and the level of human capital. In a similar fashion, a country’s ability to absorb the spillover gains from international R&D efforts
increases as the level of these key domestic variables increases. Focusing on a sample of developing countries, Coe, Helpman, and Hoffmaister (1997) show that productivity growth is significantly influenced by international R&D spillovers, where a country’s level of trade with leading-edge countries helps to determine the international transmission of R&D efforts.

For our purposes, a key contribution is provided by Aghion, Howitt, and Mayer-Foulkes (2005), who highlight the critical role played by financial development in enabling technological catch-up. These authors show that a threshold level of financial development is required if a country is to converge to the technological frontier. The underlying explanation is that domestic innovation is required in order to absorb leading-edge technologies. Such innovation is costly, and its financing requires a sufficiently developed domestic financial system. Accordingly, productivity growth is constrained if a country is not sufficiently financially developed.

Aghion and others (2009) make a related point, emphasizing the complementarity between domestic savings and foreign investment in promoting innovation. In their model, convergence to the technological frontier requires collaboration between a foreign investor and a domestic entrepreneur. The greater is the equity that the domestic agent can invest in the cooperative project, the less problematic are agency problems in operating the joint venture. In support of this hypothesis, these authors provide empirical evidence that productivity growth is positively associated with lagged savings in low-income countries. In turn, financial development is important for innovation as a result of its contribution to the mobilization of domestic savings.

In a similar fashion, Alfaro and Kalemli-Ozcan (2004) highlight the importance of domestic financial development if a developing country is to maximize the spillovers from inward foreign direct investment (FDI). In particular, these authors emphasize that domestic firms require financing if they are to reorganize production techniques in order to take advantage of the knowledge acquired through such FDI spillovers. In a similar fashion, local entrepreneurs can only imitate the technologies of foreign-owned firms if they have access to finance that would enable them to set up new operations to exploit the newly acquired knowledge. Alfaro and Kalemli-Ozcan (2004) show robust empirical evidence that the connection between FDI and economic growth is enhanced by greater domestic financial development. In subsequent work, Alfaro, Kalemli-Ozcan, and Sayek (2009) show that this connection indeed operates via total factor productivity growth. Chor, Foley, and Manova (2008) provide further evidence, showing that as the entry response of domestic firms to FDI grows higher, so also does the level of domestic financial development.

More generally, the empirical literature has established that domestic financial development is a robust correlate of faster economic growth and higher income levels (see Levine 2005 for an extensive survey). In particular, the evidence from aggregate and micro-level studies is that financial development boosts total factor productivity among advanced economies, while it also promotes growth by lowering the cost of capital in emerging and developing economies.

There are many mechanisms by which financial development may promote productivity growth, and an extensive literature investigates each channel (Demirgüç-Kunt,
Beck, and Honohan 2008 provide a detailed review). For instance, Hartmann and others (2007) emphasize the role of financial development in facilitating the reallocation of capital to faster-growing industries and find evidence in support of that channel. Hsieh and Klenow (2009) highlight the importance of resource reallocation for productivity growth in developing countries, showing that productivity growth in China and India has been driven largely by the reallocation of labor and capital from low-productivity to high-productivity firms. In a similar fashion, Song, Storesletten, and Zilibotti (2008) emphasize reallocation dynamics as a driver of Chinese productivity growth. The role of financial frictions in delaying the reallocation process is also highlighted in the quantitative model of Buera and Shin (2008).

In relation to the empirical literature on the determinants of innovation in developing countries, the World Bank has produced several significant studies in recent years. Bosch, Lederman, and Maloney (2005) examine the relation between R&D and the number of patents granted. These authors find that the effectiveness of R&D is lower in developing countries in terms of generating patentable discoveries. In turn, lower effectiveness can be related to lower levels of education and lower-quality institutional environments.

Lederman and Maloney (2003) take a broader view of the role of R&D and show that the elasticity of output growth vis-a-vis R&D is greater for developing countries, such that the social rate of return to R&D is higher for this group. However, the scale of R&D spending is much higher in higher-income countries, suggesting that there are significant barriers to R&D activity in the developing world. Their analysis identifies a low level of financial development as an important constraint on R&D activity. In addition, other country characteristics are also important, such as the protection of intellectual property rights, effectiveness of government, and quality of research institutions.

A causal connection between innovation and level of output per capita is also established in the empirical work reported by Lederman and Saenz (2005), even controlling for factors such as the quality of institutions and the level of trade openness. Accordingly, in view of the contribution of innovation activity to living standards, these authors advocate the importance of establishing the determinants of innovation activity.

Turning to firm-level evidence, Lederman (2009) studies a panel of 25,000 manufacturing firms across 68 developing and advanced economies. He finds that exporting status and the licensing of foreign technologies are good predictors of the rate of firm-level innovation. In turn, such correlates may lend further support for the hypothesis that domestic financial development is important for innovation. For instance, Manova (2008) provides extensive evidence that as exporting status is easier to attain, domestic financial development increases, since credit constraints act as a barrier to financing the fixed costs of entering export markets.

Buera, Kaboski, and Shin (2008) emphasize the connection between financial development and the expansion of the tradables sector. These authors show that production typically requires a larger scale in the tradables sector than in the non-traded sector. In the absence of financial development, the growth of the tradables sector will be constrained by the small feasible size of firms. In turn, this negatively
affects aggregate productivity growth, in view of the greater scope for high-productivity operations in the tradables sector.

Gorodnichenko, Svejnar, and Terrell (2008) also conduct a firm-level study, using data on 11,500 firms in 27 emerging-market economies. This study finds that several dimensions of globalization stimulate innovation activity. In particular, as firms are more likely to innovate, the greater is the level of foreign competition, the stronger are the vertical linkages with foreign firms, and the higher is the level of international trade. Moreover, these authors find that the gains are similar across the manufacturing and service sectors. Their finding that greater product market competition (via the entry of foreign firms) stimulates innovation activities is especially important, in view of the ambiguous effects that have been found in work focused on data for advanced economies (see, for example, Aghion and others 2006).

Aghion, Fally, and Scarpetta (2007) study the connection between financial development and the entry and growth of small firms. These authors study firm-level data for 16 advanced and emerging economies and show that access to finance promotes the entry of new firms and their growth after entry. These effects are strongest for those industrial sectors that are most dependent on external finance. Accordingly, financial development promotes the “creative destruction” process by which new technologies are disseminated through the entry of new firms.

The message from this body of work is that domestic innovation activity is required in order for a developing country to improve its level of productivity. In turn, a country’s capacity to innovate is a function of its domestic financial development, among other factors. Accordingly, by promoting financial development, international financial integration has the potential to influence positively the rate of domestic innovation activity and the rate of productivity growth.

Productivity, Innovation, and International Financial Integration

This section turns to the links between international financial integration and the levels of productivity and innovation activity.

The Real Effects of Financial Globalization

In principle, financial globalization may affect the level of productivity through several channels. Most directly, international financial integration may stimulate domestic financial development. One reason pertains to the network characteristic of financial markets: the deeper and more liquid are financial markets, the more attractive is participation in these markets because of the increasing returns to scale. Martin and Rey (2000, 2004) model this process: an increase in the scale of the financial system boosts liquidity and reduces transaction costs, which in turn increases the gains from creating new financial assets.

Levine (2001) reports evidence supporting the hypothesis that financial liberalization promotes domestic financial development. He finds that international financial integration improves the liquidity of the domestic stock market. In addition, the
efficiency of the domestic banking sector is improved by the entry of foreign-owned banks. In a more recent study, Chinn and Ito (2006) show that the contribution of financial openness to financial development is conditional on the general quality of domestic institutions. In particular, they find that the domestic equity market expands only if a threshold level of institutional quality is attained. They also find that capital account liberalization promotes financial development only if trade openness has been achieved. Finally, their empirical work underlines the complementarity between the development of the domestic banking sector and the development of equity markets.

Demirgüç-Kunt, Beck, and Honohan (2008) review the evidence on the contribution of foreign-owned banks to financial development. While the empirical literature finds quite nuanced results, these authors conclude that the balance of the evidence indicates that opening to foreign banks has the potential to increase efficiency and the level of competition. However, the gains from the entry of foreign-owned banks will be greater for those developing countries that possess the domestic institutional structures that enable foreign-owned banks to implement the higher-technology style of banking that is the comparative advantage of large, multinational banks.

In addition, international financial integration expands the scope of financial possibilities for domestic investors by virtue of the potential gains from international risk sharing. In particular, international financial integration fundamentally alters the scope for risk diversification and thereby improves access to finance for riskier projects. In turn, this may raise the trend growth rate of an economy since firms will be better able to obtain financing for higher-return, higher-risk projects. Through this mechanism, the profile for output will change to higher average growth, albeit with greater volatility (Obstfeld 1994). In relation to this mechanism, the evidence is that equity market liberalization increases the value of domestic stocks and enables domestic firms to expand capital expenditures, since a base of global investors reduces firm-specific investment risk (Henry 2000a, 2000b; Bekaert, Harvey, and Lundblad 2005; Chari and Henry 2008).

Finally, financial globalization may play a wider role in stimulating domestic institutional reforms that may in turn boost productivity through indirect mechanisms. This point is emphasized by Henry (2007) and Kose, Prasad, Rogoff, and Wei (2009). In particular, the latter describe the “ancillary” gains from financial globalization as occurring via improvements in the domestic institutional environment and enhancement of macroeconomic policy discipline. The political economy calculus that lies behind such ancillary effects has been analyzed by Rajan and Zingales (2003, 2004), among others.

**Empirical Evidence on the Link between International Financial Integration and Productivity**

Several recent studies have examined the link between financial globalization and productivity using cross-country comparative data. Bonfiglioli (2008) studies a sample of 70 countries over 1975–99 and finds that international financial integration has a robust positive effect on total factor productivity. In contrast, she finds that it has little
impact on the rate of capital accumulation, such that international financial integration primarily affects overall growth performance through the productivity channel.

Kose, Prasad, and Terrones (2009) estimate the impact of financial globalization on total factor productivity for a panel of 67 countries over 1966–2005. An important feature of this study is that it allows for a differential impact across equity-type liabilities and debt-type liabilities. This distinction turns out to be important, because the authors find that a higher level of equity-type liabilities is associated with a gain in total factor productivity. In contrast, an increase in debt-type liabilities is associated with lower total factor productivity, although that effect is attenuated in countries with better institutions and a higher level of domestic financial development.

Bekaert, Harvey, and Lundblad (2009) study a sample of 96 countries over 1980–2006. They also find a positive relation between financial globalization and total factor productivity. Moreover, they find support for the hypothesis that financial globalization affects productivity in part via its indirect impact on the level of domestic financial development, institutional quality, and macroeconomic policy discipline. Finally, they find the existence of nonlinearities, in that the impact of financial globalization depends on the initial levels of financial development and institutional quality.

In relation to firm-level evidence, Alfaro and Charlton (2006) explore the relation between international financial integration and the level of entrepreneurial activity in a country. These authors exploit a firm-level data set of approximately 24 million firms in nearly 100 countries in 1999 and 2004, such that the impact of country-level and industry-level influences on indicators of entrepreneurship can be estimated. They find robust evidence that increased international financial integration stimulates the activity of entrepreneurs along dimensions such as entry, size, and skewness of the firm-size distribution. Moreover, they find that the positive impact of international financial integration is greater in those industries that are more reliant on external finance and that entrepreneurial activity is higher in industries that have a larger share of foreign firms or that are vertically linked. Furthermore, Chari, Chen, and Dominguez (2009) find that foreign control of firms in emerging-market economies spurs technology transfer and improved governance.

Manova (2008) provides additional evidence by examining the dynamics of sectoral growth in the wake of international liberalization of equity markets. She finds that the sectors that grow most quickly are characterized by a high dependence on external-to-the-firm finance. The interpretation is that the international integration of equity markets improves the funding capacity of firms in finance-dependent sectors. Gupta and Yuan (2009) also study the sectoral impact of stock market liberalization in emerging markets and find that the greatest positive impact is on industries that are more dependent on external finance and have better opportunities for growth. A striking feature of this study is that sectoral growth is mainly due to an expansion in the size of existing firms rather than the entry of financially constrained new firms. However, the role of new-firm entry is stronger in countries that have lower regulatory barriers to entry.

Eichengreen, Gullapalli, and Panizza (2009) also consider industry-level evidence in estimating the impact of capital account liberalization. They find that capital
account liberalization disproportionately boosts the growth of finance-dependent industries. However, these gains in growth are only achieved by countries that have attained a threshold level of economic development.

In summary, the main message from the country-level, firm-level, and sectoral-level studies is that there appears to be a conditionally positive relation between international financial integration and level of productivity.

**Innovation Activity and Financial Integration: An Empirical Analysis**

This section explores whether the relation between international financial integration and level of productivity may be attributed in part to a connection between financial globalization and level of innovation activity.

**Determinants of Innovation Activity**

In order to investigate the cross-country relation between international financial integration and level of innovation activity, this paper examines the following baseline specification:

\[
INNOV_i = \alpha + \beta \cdot \log YPC_i + \gamma \cdot IFI_i + \sigma \cdot \log YPC_i \cdot IFI_i + \chi \cdot Z_i + \epsilon_i,
\]

where \(INNOV\) is an indicator of innovation activity, \(YPC\) is GDP per capita, \(IFI\) is the sum of foreign assets and foreign liabilities (expressed as a ratio to GDP), and \(Z\) is a set of general control variables. The main indicator of innovation is the ratio of R&D expenditure to GDP. However, measures of corporate dynamism are also examined, in view of the role played by firms in driving technological progress.

The level of GDP per capita is included as a general control variable. Also included is an interaction term between the \(IFI\) variable and the level of GDP per capita, in order to establish whether a threshold effect operates, by which the pattern of co-variation between financial globalization and level of innovation activity is sensitive to the level of development. The list of general control variables includes measures of general institutional quality, trade openness, educational attainment, indicators of domestic financial development, and the corporate tax rate. Since international financial integration may influence at least some of these variables, the estimated coefficients in this specification only capture the partial impact of financial globalization, while holding these characteristics fixed.

Also considered is an alternative specification allowing for a difference between equity-type and debt-type international positions. This specification has the following form:

\[
INNOV_i = \alpha + \beta \cdot \log YPC_i + \gamma_E \cdot IEQ_i + \sigma_E \cdot \log YPC_i \cdot IEQ_i + \gamma_D \cdot IDEBT_i + \sigma_D \cdot \log YPC_i \cdot IDEBT_i + \chi \cdot Z_i + \epsilon_i,
\]

where \(IEQ\) is the sum of FDI and foreign portfolio equity assets and liabilities (expressed as a ratio to GDP) and \(IDEBT\) is the sum of international debt assets and
liabilities (expressed as a ratio to GDP). The category of international debt comprises portfolio debt, non-portfolio debt, and foreign exchange reserves. As in equation 1, interaction terms are included to allow for threshold effects in the impact of different dimensions of financial globalization.

The Data

The main indicator of innovation activity is the level of expenditure on R&D, expressed as a percentage of GDP (UNESCO Innovation data set). The R&D expenditure variable captures aggregate spending on research and development but does not differentiate across different types of activity, such as the distinction between the creation of new technologies and the imitation or adaptation of existing technologies. While R&D expenditure is an “input” measure, the heterogeneity in innovation outputs (especially for developing countries) means that indicators such as number of patents may be too narrow to capture the level of innovation activity. The most recent year available is used, which is typically 2006.

In addition, Djankov and others (2009) examine a measure of business density and the business entry rate, where these variables are taken from the World Bank’s entrepreneurship survey. Alfaro and Charlton (2006) also examine firm dynamics in order to capture the role played by new enterprises and reallocation in the innovation process. The analysis here follows their lead.

The levels of foreign asset and foreign liability positions are employed in order to measure the de facto scale of international financial integration, where these data are drawn from an updated version of the data set constructed by Lane and Milesi-Ferretti (2007). The rationale for using such volume-based measures is analogous to employing the volume of exports and imports to measure the degree of effective trade openness. In particular, the impact of financial globalization on an economy should increase as the gross scale of cross-border financial positions increases.

The level of GDP per capita is measured in constant international dollars, taken from the Penn World Tables. Two standard measures of domestic financial development (the ratio of liquid liabilities to GDP and the ratio of stock market capitalization to GDP) are employed, each taken from the World Bank’s Financial Structure database. In relation to the other control variables, the measure of trade openness is taken from the World Bank’s World Development Indicators database, while the educational attainment measure (for the +25 age cohort) is taken from Barro and Lee (2001). Institutional quality is measured with the “government effectiveness” variable from the World Bank’s Governance Indicators data set. Finally, the measure of one-year effective corporate tax rate is constructed by Djankov and others (2009).

Innovation Activity and Level of Development

Figure 1 shows the cross-country relation between the level of GDP per capita and the scale of R&D expenditures (expressed as a ratio to GDP). The level of R&D spending clearly increases with the level of output per capita. Moreover, the relation is convex: the elasticity of R&D spending with respect to output per capita is
relatively small until a threshold value of output per capita is attained. After that threshold, the elasticity of R&D spending with respect to output per capita becomes substantially larger.

Two other indicators of innovation activity are considered in figures 2 and 3. Figure 2 displays the scatter plot of business density against GDP per capita, where business density is measured as the number of limited-liability corporations relative to the working-age population. Figure 3 shows the rate of business entry, which measures the number of newly registered limited-liability corporations as a ratio to the existing stock of corporations. A high level of business density and a high entry rate indicate an economic environment that is conducive to entrepreneurial activity (see also the discussion in Djankov and others 2009). Both figures show a positive connection between the level of development and the level of entrepreneurial activity, although the degree of co-variation is stronger for the business density measure than for the entry rate measure.

**Econometric Results**

The discussion now turns to an econometric analysis of the cross-country variation in levels of innovation activity, with a particular focus on the relation between international financial integration and innovation activity. Our approach is based on the specifications described in equations 1 and 2.
FIGURE 2. Relation between Business Density and Level of Development

Sources: Djankov and others 2009; Penn World Tables.

FIGURE 3. Relation between Entry Rate and Level of Development

Sources: Alfaro and Charlton 2006; Penn World Tables.
Table 1 shows the baseline results for R&D expenditure. Column 1 shows the positive relation between output per capita and level of R&D spending. Moreover, it shows a nonlinear relation between international financial integration and R&D. In particular, a higher degree of international financial integration boosts R&D spending only if a threshold level of output per capita has been attained. In fact, the estimated threshold is quite high, at 22,000 international dollars (2000 constant prices).

The basic pattern of results is very similar even when additional control variables are included in columns 2–4. In relation to these controls, trade openness is individually significant in columns 2–4: a higher level of engagement in international trade is associated with a greater level of R&D expenditure. In addition, one of the financial development variables (level of stock market capitalization) is marginally significant in column 3. Otherwise, the control variables are not individually significant.

### TABLE 1. R&D and International Financial Integration I

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)</td>
<td>–1.4</td>
<td>–2.3</td>
<td>–3.3</td>
<td>–3.2</td>
</tr>
<tr>
<td></td>
<td>(0.53)</td>
<td>(2.1)</td>
<td>(2.2)</td>
<td>(2.4)</td>
</tr>
<tr>
<td>YPC</td>
<td>0.34</td>
<td>0.24</td>
<td>0.21</td>
<td>0.19</td>
</tr>
<tr>
<td></td>
<td>(0.07)**</td>
<td>(0.1)**</td>
<td>(0.12)*</td>
<td>(0.13)</td>
</tr>
<tr>
<td>IFI</td>
<td>–0.015</td>
<td>–0.023</td>
<td>–0.027</td>
<td>–0.029</td>
</tr>
<tr>
<td></td>
<td>(0.004)**</td>
<td>(0.007)**</td>
<td>(0.008)**</td>
<td>(0.008)**</td>
</tr>
<tr>
<td>IFI*YPC</td>
<td>0.0015</td>
<td>0.0022</td>
<td>0.0026</td>
<td>0.0027</td>
</tr>
<tr>
<td></td>
<td>(0.0004)</td>
<td>(0.0006)**</td>
<td>(0.0007)**</td>
<td>(0.008)**</td>
</tr>
<tr>
<td>IQUAL</td>
<td>8.6</td>
<td>13.9</td>
<td>15.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(14.7)</td>
<td>(15.3)</td>
<td>(15.0)</td>
<td></td>
</tr>
<tr>
<td>Trade</td>
<td>0.0035</td>
<td>0.0037</td>
<td>0.0037</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0017)**</td>
<td>(0.0018)**</td>
<td>(0.0019)*</td>
<td></td>
</tr>
<tr>
<td>EDUC</td>
<td>0.009</td>
<td>0.01</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.011)</td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>LLY</td>
<td>0.24</td>
<td>0.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.34)</td>
<td>(0.33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STKCAP</td>
<td>0.31</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.17)*</td>
<td>(0.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAX</td>
<td>–0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.58</td>
<td>0.56</td>
<td>0.61</td>
<td>0.6</td>
</tr>
<tr>
<td>N</td>
<td>71</td>
<td>53</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

**Note:** Estimation is by ordinary least squares, with heteroskedasticity-consistent standard errors. YPC is GDP per capita; IFI is the ratio of foreign assets and liabilities to GDP; IQUAL is an indicator of government effectiveness; Trade is the ratio of exports plus imports to GDP; EDUC is the Barro-Lee indicator of educational attainment for the +25 age cohort; LLY is the ratio of liquid liabilities to GDP; STKCAP is the ratio of stock market capitalization to GDP; TAX is the one-year effective corporate tax rate.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.

Source: Author’s calculations.
Equation 2 is implemented in table 2 by differentiating between cross-border equity holdings and cross-border debt holdings. Column 1 shows a striking result: there is a significant relation between international equity integration and R&D expenditure, but no such relation applies for international debt integration. Moreover, the positive impact of international equity integration kicks in at a relatively low level of output per capita (3,700 international dollars).

TABLE 2. R&D and International Financial Integration II

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>-1.3</td>
<td>-2.2</td>
<td>-2.9</td>
<td>-2.8</td>
</tr>
<tr>
<td></td>
<td>(0.52)**</td>
<td>(2.1)</td>
<td>(2.2)</td>
<td>(2.40)</td>
</tr>
<tr>
<td>YPC</td>
<td>0.33</td>
<td>0.24</td>
<td>0.2</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>(0.07)***</td>
<td>(0.1)**</td>
<td>(0.12)</td>
<td>(0.13)</td>
</tr>
<tr>
<td>IEQ</td>
<td>-0.023</td>
<td>-0.02</td>
<td>-0.025</td>
<td>-0.026</td>
</tr>
<tr>
<td></td>
<td>(0.011)**</td>
<td>(0.012)</td>
<td>(0.014)*</td>
<td>(0.015)*</td>
</tr>
<tr>
<td>IEQ*YPC</td>
<td>0.0028</td>
<td>0.0024</td>
<td>0.0028</td>
<td>0.0029</td>
</tr>
<tr>
<td></td>
<td>(0.0012)**</td>
<td>(0.0014)*</td>
<td>(0.0016)*</td>
<td>(0.0017)*</td>
</tr>
<tr>
<td>IDEBT</td>
<td>-0.013</td>
<td>-0.025</td>
<td>-0.029</td>
<td>-0.03</td>
</tr>
<tr>
<td></td>
<td>(0.0027)</td>
<td>(0.01)**</td>
<td>(0.01)**</td>
<td>(0.01)***</td>
</tr>
<tr>
<td>IDEBT*YPC</td>
<td>0.0011</td>
<td>0.0022</td>
<td>0.0026</td>
<td>0.0027</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td>(0.001)**</td>
<td>(0.001)**</td>
<td>(0.001)**</td>
</tr>
<tr>
<td>IQUAL</td>
<td>8.7</td>
<td>11.9</td>
<td>13.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(14.6)</td>
<td>(14.8)</td>
<td>(14.3)</td>
<td></td>
</tr>
<tr>
<td>Trade</td>
<td>0.0014</td>
<td>0.0013</td>
<td>0.0013</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.0024)</td>
<td></td>
</tr>
<tr>
<td>EDUC</td>
<td>0.008</td>
<td>0.011</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>LLY</td>
<td>0.46</td>
<td>0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.35)</td>
<td>(0.34)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STKCAP</td>
<td>0.16</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAX</td>
<td>-0.006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.59</td>
<td>0.56</td>
<td>0.61</td>
<td>0.6</td>
</tr>
<tr>
<td>N</td>
<td>71</td>
<td>53</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

Source: Author's calculations.

Note: Estimation is by ordinary least squares, with heteroskedasticity-consistent standard errors. YPC is GDP per capita; IEQ is the ratio of foreign equity-type assets and liabilities to GDP; IDEBT is the ratio of foreign debt-type assets and liabilities to GDP; IQUAL is an indicator of government effectiveness; Trade is the ratio of exports plus imports to GDP; EDUC is the Barro-Lee indicator of educational attainment for the +25 age cohort; LLY is the ratio of liquid liabilities to GDP; STKCAP is the ratio of stock market capitalization to GDP; TAX is the one-year effective corporate tax rate.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.
The results for international equity integration broadly hold in the expanded specifications reported in columns 2–4. These specifications show a significant role for international debt integration. However, the overall association between international debt integration and R&D spending remains negative for all countries in the sample, even if it less negative for higher-output countries. Finally, none of the control variables is individually significant in columns 2–4.

Tables 3 through 6 present the results for indicators of entrepreneurial activity. Table 3 shows a very strong pattern of co-variation between GDP per capita and business density. However, columns 1–4 also show a systematic pattern in the relation between international financial integration and business density. In contrast to the R&D variable, the locus of this relation is quite different in that the positive

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>-4.7</td>
<td>-5.6</td>
<td>-5.0</td>
<td>-4.9</td>
</tr>
<tr>
<td></td>
<td>(1.1)***</td>
<td>(2.3)**</td>
<td>(2.2)**</td>
<td>(2.2)**</td>
</tr>
<tr>
<td>YPC</td>
<td>0.84</td>
<td>0.82</td>
<td>0.82</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>(0.12)***</td>
<td>(0.17)***</td>
<td>(0.19)***</td>
<td>(0.19)***</td>
</tr>
<tr>
<td>IFI</td>
<td>0.023</td>
<td>0.018</td>
<td>0.019</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>(0.007)***</td>
<td>(0.007)**</td>
<td>(0.007)***</td>
<td>(0.007)**</td>
</tr>
<tr>
<td>IFI*YPC</td>
<td>-0.0023</td>
<td>-0.0018</td>
<td>-0.0018</td>
<td>-0.0018</td>
</tr>
<tr>
<td></td>
<td>(0.0007)***</td>
<td>(0.0006)***</td>
<td>(0.0006)***</td>
<td>(0.007)**</td>
</tr>
<tr>
<td>IQ( \text{QUAL} )</td>
<td>7.2</td>
<td>5.8</td>
<td>7.2</td>
<td>7.2</td>
</tr>
<tr>
<td></td>
<td>(11.8)</td>
<td>(12.1)</td>
<td>(13.8)</td>
<td></td>
</tr>
<tr>
<td>Trade</td>
<td>-0.0004</td>
<td>-0.001</td>
<td>-0.001</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td></td>
</tr>
<tr>
<td>EDUC</td>
<td>-0.0018</td>
<td>-0.011</td>
<td>-0.012</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td></td>
</tr>
<tr>
<td>LLY</td>
<td>-0.18</td>
<td>-0.13</td>
<td>-0.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.35)</td>
<td>(0.39)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STK( \text{CAP} )</td>
<td>0.07</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>(0.43)</td>
<td>(0.17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAX</td>
<td>-0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.53</td>
<td>0.59</td>
<td>0.61</td>
<td>0.6</td>
</tr>
<tr>
<td>N</td>
<td>60</td>
<td>44</td>
<td>43</td>
<td>43</td>
</tr>
</tbody>
</table>

Source: Author's calculations.

Note: Estimation is by ordinary least squares, with heteroskedasticity-consistent standard errors. YPC is GDP per capita; IFI is the ratio of foreign assets and liabilities to GDP; IQ\( \text{QUAL} \) is an indicator of government effectiveness; Trade is the ratio of exports plus imports to GDP; EDUC is the Barro-Lee indicator of educational attainment for the +25 age cohort; LLY is the ratio of liquid liabilities to GDP; STK\( \text{CAP} \) is the ratio of stock market capitalization to GDP; TAX is the one-year effective corporate tax rate.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.
association between international financial integration and business density is stronger at lower levels of output per capita. Indeed, it turns negative at higher levels of output per capita (beyond 22,000 international dollars). None of the control variables is significant in columns 2–4.

Table 4 distinguishes between international equity integration and international debt integration. The results show that international equity integration is significantly positively associated with a higher level of business density. Moreover, the

| TABLE 4. Business Density and International Financial Integration II |
|----------------|----------|----------|----------|--------|
| Variable       | (1)      | (2)      | (3)      | (4)    |
| α              | –4.3     | –5.5     | –4.2     | –4.2   |
| (1.1)***       | (2.2)**  | (2.1)*   | (2.1)*   |
| YPC            | 0.8      | 0.8      | 0.78     | 0.78   |
| (0.12)***      | (0.18)***| (0.19)***| (0.19)   |
| IEQ            | 0.053    | 0.073    | 0.043    | 0.043  |
| (0.03)*        | (0.036)* | (0.026)  | (0.026)  |
| IEQ*YPC        | –0.005   | –0.007   | –0.0038  | –0.0038|
| (0.0032)       | (0.0037)*| (0.0028) | (0.0028) |
| IDEBT          | 0.0019   | –0.024   | –0.0024  | –0.0026|
| (0.019)        | (0.03)   | (0.02)   | (0.02)   |
| IDEBT*YPC      | –0.0003  | –0.0023  | 0.0001   | 0.0002 |
| (0.002)        | (0.003)  | (0.002)  | (0.002)  |
| IQUAL          | 10.1     | 4.4      | 4.9      |        |
| (11.3)         | (12.0)   | (13.8)   |          |
| Trade          | –0.002   | –0.0036  | –0.0036  |        |
| (0.003)        | (0.0027) | (0.0028) |          |
| EDUC           | –0.014   | –0.015   | –0.016   |        |
| LLY            |          | 0.048    | 0.06     |        |
| (0.38)         | (0.42)   |          |          |
| STKCAP         |          | –0.14    | –0.14    |        |
| (0.19)         | (0.19)   |          |          |
| TAX            |          |          | –0.003   |        |
| (0.03)         |          |          |          |
| R²             | 0.54     | 0.62     | 0.62     | 0.61   |
| N              | 60       | 44       | 43       | 43     |

Source: Author’s calculations.

Note: Estimation is by ordinary least squares, with heteroskedasticity-consistent standard errors. YPC is GDP per capita; IEQ is the ratio of foreign equity-type assets and liabilities to GDP; IDEBT is the ratio of foreign debt-type assets and liabilities to GDP; IQUAL is an indicator of government effectiveness; Trade is the ratio of exports plus imports to GDP; EDUC is the Barro-Lee indicator of educational attainment for the +25 age cohort; LLY is the ratio of liquid liabilities to GDP; STKCAP is the ratio of stock market capitalization to GDP; TAX is the one-year effective corporate tax rate.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.
interaction term is only marginally significant in column 2 and, even then, the overall association between international equity integration and business density only turns negative at a very high level of output (33,400 international dollars). In relation to international debt integration, there is no significant association between cross-border debt holdings and level of business density. As in table 3, none of the control variables is individually significant.

Finally, tables 5 and 6 present the findings for rate of business entry. Consistent with the scatter plot in figure 3, there is a positive but mild relation between GDP per capita and the business entry rate. Columns 2–4 provide some evidence of co-variation between international financial integration and the business entry rate. As with the R&D measure, the pattern is negative until a threshold level of output

### TABLE 5. Entry Rate and International Financial Integration I

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>−3.4</td>
<td>−3.6</td>
<td>−3.8</td>
<td>−3.8</td>
</tr>
<tr>
<td></td>
<td>(0.5)***</td>
<td>(1.0)***</td>
<td>(1.1)***</td>
<td>(1.2)***</td>
</tr>
<tr>
<td>YPC</td>
<td>0.09</td>
<td>0.15</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>(0.056)*</td>
<td>(0.062)**</td>
<td>(0.055)***</td>
<td>(0.056)***</td>
</tr>
<tr>
<td>IFI</td>
<td>0.002</td>
<td>−0.009</td>
<td>−0.007</td>
<td>−0.007</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.005)</td>
<td>(0.004)*</td>
<td>(0.004)*</td>
</tr>
<tr>
<td>IFI*YPC</td>
<td>−0.0002</td>
<td>0.0008</td>
<td>0.0008</td>
<td>0.0008</td>
</tr>
<tr>
<td></td>
<td>(0.0004)</td>
<td>(0.0005)*</td>
<td>(0.0004)*</td>
<td>(0.0004)*</td>
</tr>
<tr>
<td>QUAL</td>
<td>1.4</td>
<td>2.3</td>
<td>2.4</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>(5.5)</td>
<td>(6.4)</td>
<td>(6.7)</td>
<td></td>
</tr>
<tr>
<td>Trade</td>
<td>0.0022</td>
<td>0.0015</td>
<td>0.0015</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0012)*</td>
<td>(0.0013)</td>
<td>(0.0013)</td>
<td></td>
</tr>
<tr>
<td>EDUC</td>
<td>−0.014</td>
<td>−0.021</td>
<td>−0.021</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)**</td>
<td>(0.005)***</td>
<td>(0.005)***</td>
<td></td>
</tr>
<tr>
<td>LLY</td>
<td>−0.62</td>
<td>−0.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.15)***</td>
<td>(0.16)***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STKCAP</td>
<td>0.009</td>
<td>0.008</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td>(0.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAX</td>
<td>−0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.13</td>
<td>0.25</td>
<td>0.42</td>
<td>0.4</td>
</tr>
<tr>
<td>N</td>
<td>57</td>
<td>43</td>
<td>42</td>
<td>42</td>
</tr>
</tbody>
</table>

Source: Author’s calculations.

Note: Estimation is by ordinary least squares, with heteroskedasticity-consistent standard errors. YPC is GDP per capita; IFI is the ratio of foreign assets and liabilities to GDP; QUAL is an indicator of government effectiveness; Trade is the ratio of exports plus imports to GDP; EDUC is the Barro-Lee indicator of educational attainment for the +25 age cohort; LLY is the ratio of liquid liabilities to GDP; STKCAP is the ratio of stock market capitalization to GDP; TAX is the one-year effective corporate tax rate.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.
per capita is attained. However, that threshold level is relatively low: $6,300 in the specification reported in column 3.

In terms of the control variables, the educational attainment variable enters with a significantly negative sign across columns 2–4. One of the financial development measures (the ratio of liquid liabilities to GDP) is also significantly negative in columns 3–4. However, the trade variable is positive and marginally significant in column 2.

Table 6 shows that international equity integration and international debt integration are both significant across columns 2–4. Indeed, the pattern is quite striking.

### Table 6. Entry Rate and International Financial Integration II

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>-3.3</td>
<td>-3.5</td>
<td>-3.9</td>
<td>-3.9</td>
</tr>
<tr>
<td></td>
<td>(0.5)***</td>
<td>(1.1)***</td>
<td>(1.3)***</td>
<td>(1.3)***</td>
</tr>
<tr>
<td>YPC</td>
<td>0.077</td>
<td>0.14</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.058)**</td>
<td>(0.05)***</td>
<td>(0.05)***</td>
</tr>
<tr>
<td>IEQ</td>
<td>0.019</td>
<td>0.036</td>
<td>0.027</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.014)**</td>
<td>(0.011)**</td>
<td>(0.011)**</td>
</tr>
<tr>
<td>IEQ*YPC</td>
<td>-0.002</td>
<td>-0.0038</td>
<td>-0.0029</td>
<td>-0.0029</td>
</tr>
<tr>
<td></td>
<td>(0.0014)</td>
<td>(0.0015)**</td>
<td>(0.0012)**</td>
<td>(0.0012)**</td>
</tr>
<tr>
<td>IDEBT</td>
<td>-0.008</td>
<td>-0.041</td>
<td>-0.032</td>
<td>-0.032</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.012)***</td>
<td>(0.008)***</td>
<td>(0.008)***</td>
</tr>
<tr>
<td>IDEBT*YPC</td>
<td>0.001</td>
<td>0.0042</td>
<td>0.0033</td>
<td>0.0033</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.0012)***</td>
<td>(0.0008)***</td>
<td>(0.0008)***</td>
</tr>
<tr>
<td>IQUAL</td>
<td>3.3</td>
<td>3.7</td>
<td>4.0</td>
<td>-7.8</td>
</tr>
<tr>
<td></td>
<td>(6.2)</td>
<td>(7.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade</td>
<td>0.0027</td>
<td>0.0026</td>
<td>0.0026</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0014)*</td>
<td>(0.0008)***</td>
<td>(0.0008)***</td>
<td></td>
</tr>
<tr>
<td>EDUC</td>
<td>-0.021</td>
<td>-0.025</td>
<td>-0.026</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)***</td>
<td>(0.005)***</td>
<td>(0.005)***</td>
<td></td>
</tr>
<tr>
<td>LLY</td>
<td>-0.61</td>
<td>-0.61</td>
<td>-0.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.17)***</td>
<td>(0.17)***</td>
<td>(0.17)***</td>
<td></td>
</tr>
<tr>
<td>STKCAP</td>
<td>-0.0033</td>
<td>-0.0036</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAX</td>
<td></td>
<td></td>
<td>-0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.14</td>
<td>0.39</td>
<td>0.51</td>
<td>0.49</td>
</tr>
<tr>
<td>N</td>
<td>57</td>
<td>43</td>
<td>42</td>
<td>42</td>
</tr>
</tbody>
</table>

Source: Author's calculations.

Note: Estimation is by ordinary least squares, with heteroskedasticity-consistent standard errors. YPC is GDP per capita; IEQ is the ratio of foreign equity-type assets and liabilities to GDP; IDEBT is the ratio of foreign debt-type assets and liabilities to GDP; IQUAL is an indicator of government effectiveness; Trade is the ratio of exports plus imports to GDP; EDUC is the Barro-Lee indicator of educational attainment for the +25 age cohort; LLY is the ratio of liquid liabilities to GDP; STKCAP is the ratio of stock market capitalization to GDP; TAX is the one-year effective corporate tax rate.

*** Significant at the 1 percent level.
** Significant at the 5 percent level.
* Significant at the 10 percent level.
was the case for the business density variable, international equity integration is associated with a higher entry rate at lower levels of GDP per capita, but the relation turns negative beyond a threshold level. In fact, the threshold level is lower than was the case for business density, kicking in at $13,000.

In contrast, international debt integration has a negative association with the entry rate at lower levels of output per capita. However, the relation turns positive once output per capita exceeds $17,300. Taken together, these results suggest that international equity integration is especially useful at lower levels of development, whereas international debt integration can be helpful, but only at higher levels of development. The results for the control variables are quite similar to those for the business density variable. The main exception is that the trade variable is now significant across columns 2–4 and is highly significant in columns 3 and 4.

Of course, these empirical results are highly provisional. While the partial correlations uncovered by the regression analysis are intriguing, much more extensive research is required in order to probe the links between financial globalization and innovation activity in a more complete manner. In one direction, it is well understood that results in cross-country cross-sectional regression studies can be sensitive to the choice of control variables. Accordingly, it will be important in future research to investigate alternative sets of control variables.

This paper has not attempted to establish whether the links between international financial integration and innovation activity are causal in nature. If a good set of instruments for international financial integration could be identified, it would be useful to explore endogeneity issues.

In relation to the role of threshold effects, this paper has opted to focus on the level of GDP per capita as the variable that influences the relation between international financial integration and innovation activity. This has the virtue of being a fairly general indicator of the level of development. However, other authors have sought to examine threshold effects where specific institutional variables or other structural characteristics (such as the level of domestic financial development) intermediate the relation between international financial integration and various macroeconomic outcomes (see Kose, Ayhan, Prasad, and Taylor 2009 for a recent synthesis). Accordingly, a challenge for future research is to establish with a greater level of precision the exact threshold conditions that are required in order to reap the gains from international financial integration. Moreover, it seems clear that there are multiple thresholds, with the threshold conditions for international equity integration less demanding than those for international debt integration.

Further research should include time-series analysis, in order to establish under what circumstances countries that increase the level of international financial integration also experience an increase in innovation activity. However, such time-series analysis is hampered by relatively short time series for some of the key variables. Moreover, establishing the line of causation in time-series studies can be especially difficult in view of the timing issues involved. Finally, the literature on financial development in emerging-market economies and developing countries has found that episodes of major financial liberalization frequently involve a crisis phase in
which excess debt levels lead to banking and currency crises. Rancière, Tornell, and
Westermann (2008) present evidence that liberalization still raises long-term
growth, even accounting for such “bumpiness.” Since innovation activity varies
over the business cycle, a further task for future research is to establish the impact
of financial crises on the level of innovation activity.

The evidence that greater financial integration is associated with higher levels of
innovation activity (conditional on threshold levels of development) helps to support
the long-term case in favor of financial globalization. However, this is not sufficient
to suggest that a rapid move toward external financial liberalization is desirable for
all countries. Rather, in line with the cumulative evidence on associated topics in the
recent empirical literature, the existence of threshold effects means that there may be
a complex interplay between financial integration and level of innovation activity.
The gains from international equity integration seem to be more attainable for lower-
income countries than is the case for international debt integration, which suggests
that a strategy of opening up to equity-type cross-border flows should precede the
liberalization of debt-type flows. Accordingly, a country that does not have in place
the required domestic institutional and policy environment to gain fully from finan-
cial integration would be better advised to pursue financial integration in an incre-
mental fashion (see also Obstfeld 2009).

That said, it is also possible that financial integration can contribute to a better
domestic institutional framework, such that it is not necessarily appropriate to delay
all moves toward financial openness until the domestic institutional framework is
fully in place. This is in line with the research that attempts to link the political econ-
omy of reform with the degree of financial openness (see, among others, Rajan and
Zingales 2003, 2004; Gourinchas and Jeanne 2005; Mishkin 2006, 2009). Accord-
ingly, the negotiation and modulation of the appropriate pace of financial integration
remains a major challenge for policy makers in developing countries.

Moreover, it is important to emphasize that the relation between international
financial integration and national economic performance also depends on the
nature of the global governance of the international financial system. In particular,
the recent reforms of the International Monetary Fund and the expanded role for
the Financial Stability Board should help to reduce the risk profile of internation-
al financial flows for emerging-market economies (see also Lane 2009a, 2009b).
For example, the establishment of the International Monetary Fund’s flexible cred-
ity line facility is intended to reduce the risk of a well-behaved emerging-market
economy encountering liquidity problems due to external dysfunction in interna-
tional financial markets. Further governance reforms along these lines may alter
the calculus for developing countries in determining the pace of international
financial liberalization.

Conclusions

This paper has examined the connection between international financial integration
and enhanced levels of productivity and innovation activity. At a time when the
gains to financial globalization are being called into question, it is important to examine the potential long-term gains from financial integration. The evidence in this paper, together with the findings in the related recent literature, suggests that financial globalization can conditionally raise the level of innovation activity, which in turn boosts long-term productivity.

However, more research is required in order to establish whether this finding is robust and involves a causal mechanism running from financial globalization to more intense levels of innovation activity. In addition, the cross-sectional econometrics deployed in this paper should be supplemented by alternative empirical approaches, including detailed country studies that record how financial integration has affected innovation activity in specific settings. A further avenue for future research is to establish why different types of financial integration (equity positions versus debt positions) have different patterns of co-variation with innovation activity.

The challenge for policy makers in developing countries is to embrace financial globalization in a phased way that recognizes the interplay between domestic institutional development and greater openness to international investment flows. In terms of sequencing, this paper and other recent contributions find evidence that international equity integration offers greater benefits for lower-income countries than international debt integration, where the gains are concentrated at higher income levels. Finally, the burden on national policy makers can be alleviated by internationally coordinated actions to improve the stability of the global financial system. Accordingly, the G-20 reform agenda can make it safer for developing countries to obtain the potential benefits from financial globalization.

Notes

1. This evidence is reviewed later in the paper.
2. The macroeconomics of innovation has been studied extensively over the last two decades, with major overviews provided by Grossman and Helpman (1991), Aghion and Howitt (1998, 2009), and Acemoglu (2008).
3. See also Ciccone and Papaioannou (2009) on human capital and innovation.

References


Innovation, Competitiveness, and Growth: Korean Experiences

SUNGCHUL CHUNG

Science, technology, and innovation have been cited as one of the key factors behind the economic success of the Republic of Korea. By making continuous and massive investments in research and development and in innovation, Korea has succeeded in building a unique innovation system that supports sustainable growth of the Korean economy. The factors that have influenced the Korean innovation system the most are (1) outward-looking development strategy, (2) large firm–oriented industrial policy, and (3) human resources, among many others. These are the sources of both the strength and the weakness of the system. This paper reviews the developments of science and technology in Korea, evaluates their impacts on industrial development, and attempts to derive from the discussions some lessons that may be applicable to latecomers.

Over the course of four decades, the Republic of Korea has transformed itself from a stagnant agrarian society into one of the world’s most dynamic industrial economies. In the early 1960s, when Korea first launched its industrialization efforts, it was a typical developing country with a poor base of resources and production, a small domestic market, and a large population. Korea’s gross national product (GNP) in 1962 was only $2.3 billion (in 1980 prices) or $87 per capita, which came mainly from the primary sectors. The manufacturing sector’s share of GNP remained at a mere 15 percent. International trade was also at a very infant stage: in 1962 the volume of exports was only $55 million, and the volume of imports was $390 million. But Korea is now the thirteenth largest economy and one...
of the major trading countries in the world. It has achieved world prominence in areas such as semiconductors, liquid crystal displays, telecommunications equipment, automobiles, shipbuilding, and so on. Indeed, it is one of the key players in the global economy.

Korea has achieved in four decades what it took more than a century for the Western industrial countries to accomplish. A rich literature on Korean growth attributes Korea’s success to an assortment of factors, but there is broad agreement that the Korean government’s “outward-looking development strategy,” well-educated and well-disciplined workforce, and technological innovation have combined to bring about what is called the “Korean miracle.”

Of the three factors, this paper focuses on the role of technological innovation. What stimulated and facilitated Korean industries to engage so actively in research and development (R&D) and innovation? What has been the role of government in the process? And, to what extent has technological innovation improved the competitiveness of Korean industries and supported economic growth? The paper discusses these issues, with a view to drawing some lessons from the Korean development experience. The paper starts by reviewing what Korea has done to learn and acquire technologies for industrialization and how Korea has promoted R&D and innovation and built up technological capability. It then analyzes the contribution of R&D and innovation to industrial competitiveness and economic growth in Korea. Finally, it draws some lessons for latecomers.

How Korea Learned and Acquired Technologies for Industrialization

In the 1960s Korea was barren in the fields of science and technology. There were only two public institutes for scientific research and technological development: the National Defense R&D Institute, which was created right after the end of the Korean War, and the Korea Atomic Energy Research Institute, which was founded in 1959; there were fewer than 5,000 research scientists and engineers in the public and private sectors combined.2 In 1963 R&D expenditures remained at $9.5 million.

Fortunately, Korea had a well-educated workforce relative to other developing countries. The Korean workforce had an average of 4.98 years of schooling in 1960, and the elementary school enrollment rate reached 100 percent as early as 1970. Korea’s educational attainment in the 1960s stood fairly close to the level expected for a country twice as wealthy as Korea (Cohen and Soto 2001). For this, Korea owes much to the Confucian tradition, which holds education and scholarship in high esteem.3

In this setting, Korea launched the First Five-Year Economic Development Plan in 1962. Lacking technological capability, Korea had to rely almost totally on foreign sources of technology. Korea’s policy strategy was geared to promoting the inward transfer of foreign technologies, while, at the same time, developing domestic capacity to digest, assimilate, and improve upon the transferred technologies.
Korean Strategy for Technology Learning

Foreign direct investment (FDI) is often cited as a key to technological learning and one of the most effective means for latecomers to learn new production skills and acquire managerial expertise. However, the Korean government discouraged FDI by restricting ownership and repatriation of profits and imposing requirements on technology transfer and exports. Such a restrictive policy was inevitable because the public viewed multinationals as perpetuating the country’s economic and technological dependence and as reinforcing the asymmetrical relationship between the industrial and developing countries (Koo 1986; Vernon 1977; Stewart 1978).4 For this reason, FDI played a less important role in Korea’s acquisition of capital and technology than it did in other developing countries.

The purchase of technology through foreign licensing was also of modest importance in Korea because of the government’s imposition of foreign exchange controls. Being a typical agrarian economy relying on agriculture for almost three-quarters of national production, Korea in the 1960s simply could not afford to buy technology from foreign sources. It therefore curtailed foreign licensing, which often entails long-term financial commitments.

As an alternative to foreign licensing, Korea financed industrial investments through long-term foreign loans. The Korean government brought in large-scale foreign loans and allocated them to investments in select industries, leading to massive importation of foreign capital goods and turnkey plants (see table 1). Industries later reverse-engineered the imported capital goods for the purpose of acquiring needed technologies. The government selected not only the target industries for investment but also the entrepreneurs who would implement the new investment projects, and some of those entrepreneurs later became the owners of “chaebols” (defined and discussed later in this paper). The FDI policy had much to do with the unique industrial structure of Korea.

How Private Industries Responded

The response of private companies to such restrictive policies varied across industries. In the case of light industries, such as shoes, clothing, textiles, and some intermediate goods for import substitution as well as export, the major sources of technological learning were OEM (original equipment manufacturing) production arrangements.

### TABLE 1. Channels for Technology Transfer in Korea, 1962–81

<table>
<thead>
<tr>
<th>Time period</th>
<th>FDI</th>
<th>Foreign licensing</th>
<th>Capital goods</th>
</tr>
</thead>
<tbody>
<tr>
<td>1962–66</td>
<td>45.4</td>
<td>0.8</td>
<td>316.0</td>
</tr>
<tr>
<td>1967–71</td>
<td>218.6</td>
<td>16.3</td>
<td>2,541.0</td>
</tr>
<tr>
<td>1972–76</td>
<td>879.4</td>
<td>96.6</td>
<td>8,841.0</td>
</tr>
<tr>
<td>1977–81</td>
<td>720.6</td>
<td>451.4</td>
<td>27,978.0</td>
</tr>
</tbody>
</table>

Source: Data from the National Statistical Office.
Korean firms benefited most from such arrangements because they offered opportunities to work with foreign buyers who provided everything from product designs and materials to quality control at the end of the process. This was especially so in the case of garment and electronic industries (Hobday 1995).

In the 1970s Korea’s development target shifted to more capital- and technology-intensive industries, and the government implemented massive investment projects to build up machinery and chemical industries. For the development of chemical industries, Korea relied largely on the importation of turnkey plants, which offered technical training programs as part of the package. In the case of heavy machinery, foreign licensing was an important channel for technology acquisition (Chung and Branscomb 1996).

To help industries to adopt new technologies, the government created government R&D institutes in the fields of heavy machinery and chemicals, such as the Korea Institute of Machinery and Metals, the Electronics and Telecommunications Research Institute, the Korea Research Institute of Chemical Technology, the Korea Research Institute of Standards and Science, the Korea Institute for Energy Research, and the Korea Ocean R&D Institute. These institutes worked with private industries to build a technological foundation for industrial development.

As a result, FDI had a minimal impact on the Korean economy, accounting for only 4 percent of Korea’s cumulative total long-term foreign capital over the period of 1962–82 ($9 billion). According to a United Nations report, FDI in all developing countries in the early and mid-1970s accounted for 10–20 percent of their total foreign capital inflow (Ahn 1991). Over the period of 1962–71, FDI inflow in Korea remained at $264 million, while imported capital goods reached $2.9 billion.

In short, Korean industries acquired technology more from informal than formal channels. As informal channels involve less market mediation, they are less costly; they also require recipients to have higher capacity, not just in identifying and selecting technologies, but also in absorbing, assimilating, and improving upon the transferred technologies. Korea was able to acquire technologies for industrialization through informal channels as a result of its rich pool of well-educated, motivated people.

**Building a Base for R&D and Innovation**

While promoting technological learning for industrialization, the government also made efforts to build a base for science and technology (S&T) development during this period. The Science and Technology Promotion Act and the Science Education Act were passed in 1967 as a legal base for S&T development. The Korea Institute of Science and Technology (KIST) was established in 1966, followed in 1967 by the Ministry of Science and Technology (MOST), the central government agency responsible for S&T policy. In 1970 the government enacted the Korea Advanced Institute of Sciences Act, which created the basis for the Korea Advanced Institute of Sciences (KAIS; currently KAIST). KIST was the first organization dedicated strictly to R&D in Korea, while KAIS brought the U.S. graduate education system to Korea. In the 1970s various government research and development institutes were also established.
to assist industries in absorbing and assimilating technologies. In the early stage of development, these institutions made two important contributions: first, they helped industries to acquire new technologies, and second, they helped to build indigenous R&D capability by repatriating many established scientists and engineers from abroad. In 1974 the government started construction of the Daeduk Science Town, where many public and private R&D institutes, including the government research and development institutes, are now clustered.

Costs and Benefits of the Korean Strategy

The industrialization of Korea in the early phase was a process of learning how to absorb and improve on imported foreign technologies for industrial development. Technological learning, as opposed to indigenous technology development, was at the core of the development strategy. These efforts brought positive results. Around the end of the 1970s, Korea began to export items such as ships, semiconductors, and television sets.

The Korean approach to acquiring technology had both positive and negative effects. On the positive side, this policy enabled Korea to acquire technologies at lower cost and precluded the constraints often imposed by multinationals on the efforts of local firms to develop their own capability. The approach was effective in maintaining independence from the dominance of multinationals. On the negative side, Korea had to give up access to new technologies that might have been available through direct equity links with foreign firms. By restricting FDI, Korea failed to adopt global standards in domestic business operations. The most important lesson here is that had it not been for the well-educated workforce, it would not have been possible for Korea to succeed in acquiring and using technologies through informal modes of technology transfer.

How Korea Has Been Able to Build Up Indigenous R&D Capability

As industrial development continued into the 1980s, the technological requirements of Korean industries became more complex and sophisticated, making it increasingly difficult for Korea to acquire technologies through informal channels. Even though the government eased its restrictions on FDI and liberalized foreign licensing, FDI and foreign licensing did not increase significantly. To sustain development, many saw the need to build indigenous R&D capability.

From Technology Learning to Technology Development: A Policy Shift

The government responded to the changes in the technology environment by launching the National R&D Program in 1982 and taking various policy measures to promote and facilitate private R&D activities, including tax incentives, financial assistance, procurement, and other promotional actions.
First, the government introduced policy incentives that provide private firms with tax exemptions or tax breaks for investments in R&D and human resource development. In 1981 the government revised the Local Tax Law and the National Tax Reduction Law to exempt real estate dedicated to R&D purposes from local taxes and reduce the corporate tax for expenditures on R&D, human resource development, and related facilities. One year after that, the Tariff Law was changed to reduce tariffs on imported materials and instruments for R&D.

Second, in 1981 the government created the Korea Technology Development Corporation, a bank specializing in technology financing, in order to facilitate the development and commercialization of technology. The government reinforced financial supports for technology development by establishing a nonprofit guarantee institution called the Korea Technology Finance Corporation in 1989, whose major function is to help small and medium enterprises (SMEs) to obtain loans for development or commercialization of technology. The government complemented these measures by establishing the Industrial Development Fund (1986), the Science and Technology Promotion Fund (1991), the Information and Telecommunication Technology Fund (1993), and other programs to facilitate commercialization of new technologies. The financial system for technology development in Korea was completed with the opening of KOSDAQ, a technology stock market.

Third, a procurement program was introduced in 1981 to promote demand for new technology products developed by SMEs, followed by many other support programs, such as technical and legal consultancy services, technology information services, technology trade and transfer, and so on.

Fourth, government R&D investments have increased significantly since the early 1980s, when Korea’s S&T policy underwent a fundamental shift from technology learning to technology development. The Ministry of Science and Technology launched the National R&D Program in 1982, which was followed by the Industrial Base Technology Development Program of the Ministry of Commerce, Industry, and Energy, the Information and Telecommunication Technology Development Program of the Ministry of Information and Communication, and others. These initiatives promoted industrial R&D by providing private industries with opportunities to cooperate with the public sector and by inducing private R&D investments for commercial purposes.

According to a study conducted by the Science and Technology Policy Institute (STEPI 2005), the Korean government was offering 259 programs in support of industrial R&D and innovation. R&D subsidies accounted for 30 percent of all support programs, technology transfer accounted for 13 percent, and support for human resources development accounted for 11 percent. The government expended $3.4 billion for loan programs, and $3.3 billion for R&D subsidy programs. The tax revenue forgone because of the tax incentive programs amounted to $1.5 billion in 2005. In contrast, human resource development support programs received only 1 percent of the total budget spent that year (see table 2).

Figure 1 describes the R&D support policy of the Korean government. As indicated, Korea’s S&T policy focused on promoting learning from foreign sources as well as building up infrastructure for R&D and human resource development before
1980. After 1980, it shifted toward nurturing indigenous R&D capability and, at the same time, promoting and facilitating private industrial R&D.

This marked a turning point for the Korean innovation system not only because of the changes in policy orientation but also because of the changes in the role of government in science, technology, and innovation. Before the change, the government set the development target, selected the strategic technologies, and financed the implementation of development programs. So, in the early development process, the government played the role not only of planner and rule setter but also of financier. This was particularly true in the areas of R&D and innovation.7

### TABLE 2. Industrial R&D and Innovation Support Programs in Korea, 2005

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of programs</th>
<th>% of all programs</th>
<th>Budget US$ millions</th>
<th>% of all program budgets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tax incentives</td>
<td>17</td>
<td>6.6</td>
<td>1,480(^a)</td>
<td>15.9</td>
</tr>
<tr>
<td>Loans</td>
<td>15</td>
<td>5.8</td>
<td>3,402(^b)</td>
<td>36.6</td>
</tr>
<tr>
<td>Procurement</td>
<td>2</td>
<td>0.8</td>
<td>394</td>
<td>4.2</td>
</tr>
<tr>
<td>Human resources development</td>
<td>29</td>
<td>11.2</td>
<td>106</td>
<td>1.1</td>
</tr>
<tr>
<td>R&amp;D subsidies</td>
<td>77</td>
<td>29.7</td>
<td>3,253</td>
<td>35.0</td>
</tr>
<tr>
<td>Technology trade</td>
<td>8</td>
<td>3.1</td>
<td>61</td>
<td>0.7</td>
</tr>
<tr>
<td>Technology transfer</td>
<td>33</td>
<td>12.7</td>
<td>225</td>
<td>2.4</td>
</tr>
<tr>
<td>Technology consultancy</td>
<td>27</td>
<td>10.4</td>
<td>44</td>
<td>0.5</td>
</tr>
<tr>
<td>Legal assistance</td>
<td>29</td>
<td>11.2</td>
<td>34</td>
<td>0.4</td>
</tr>
<tr>
<td>Technology information</td>
<td>22</td>
<td>8.5</td>
<td>294</td>
<td>3.2</td>
</tr>
<tr>
<td>Total</td>
<td>259</td>
<td>100.0</td>
<td>9,296</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: Shin and others 2006a.

\(^a\) Tax revenues forgone.

\(^b\) Amount of loan available.

### FIGURE 1.
Development of S&T Policy in Korea

Notes: MOST = Ministry of Science and Technology; HAN = Highly Advanced National R&D Program; KOSDAQ = a technology stock market.
How Industries Responded to the Change

Private industries responded to the changes in policy by investing massive amounts in R&D. Most of all, the number of industrial R&D centers registered at the Korea Industrial Technology Association grew very rapidly: there were 129 industrial R&D centers in 1984, 200 in 1986, 2,000 in 1995, and more than 17,000 in 2009. Consequently, the relationship between technology imports and R&D changed. The ratio of technology imports to business expenditures on R&D declined sharply, from more than 90 percent in 1975 to 30 percent in the mid-1980s.8

As figure 2 shows, in the 1970s and early 1980s, the ratio of royalty payments to business expenditures on R&D in Korea was almost 100 percent, which means that Korean industries spent as much money on licensing foreign technologies as on R&D. But the ratio declined to 30 percent in the mid-1980s and to less than 20 percent in the early 1990s. In contrast, the ratio of business expenditures on R&D to sales rose rapidly, from a mere 0.5 percent in the 1970s and the early 1980s to more than 2 percent in the 1990s. By the early 1990s, Korean industries’ mode of technology acquisition shifted from borrowing and learning from foreign sources to conducting indigenous R&D.

R&D investment has since undergone a quantum leap. Korea’s R&D investment, which stood at only W 368.8 billion ($526 million, 0.81 percent GDP) in 1981, rose steadily to reach W 31.3 trillion ($33.7 billion, 3.47 percent of GDP) in 2007. Over a period of 25 years, investment in R&D increased more than 60 times. Korea invests a larger share of its income in R&D than other countries with the same or higher income. Korea now is the sixth largest spender on R&D among Organisation for Economic Co-operation and Development (OECD) countries.

FIGURE 2.
Changing Relationship between R&D and Technology Borrowing in Korea, 1976–2002

R&D Structure

The rapid growth of R&D in Korea has been led by the private sector. Today, private industries account for about 75 percent of the nation’s gross R&D expenditures, which means that private industries are the dominant players in Korea’s R&D. Of the total industrial R&D expenditures in 2007, manufacturing industries took up about 90 percent. Manufacturing R&D is largely led by electronic equipment (including office equipment, electronic components, and communications equipment, at 49 percent), automobiles (18 percent), and chemicals (11 percent). Electronic equipment, automobiles, and chemical industries absorb almost 80 percent of manufacturing R&D, suggesting a very high concentration of manufacturing R&D in a few industries.

The average R&D intensity of Korean industries was 2 percent in 2007, and the intensity of manufacturing industries was 3 percent. The industries that invested the largest share of sales in R&D include medical and precision equipment (8 percent), communications equipment (7 percent), and electronic components (6 percent). In contrast, construction, pulp and paper, food and beverages, and textile industries invested less than 1 percent of sales in R&D (see table 3).

Industrial research is highly focused on development (72 percent) and applied research (13 percent), with the remaining 13 percent devoted to basic research. Of the total industrial R&D, 80 percent was for new product development, while only 20 percent was for the improvement or development of processes.

Industrial R&D activities in Korea are highly concentrated in large enterprises. The 20 largest companies account for 56 percent of the total manufacturing R&D, the top 10 companies account for 50 percent, and the top five companies account for 44 percent. In electronic components, the share of the top 20 companies is 91 percent (see table 4). The extremely high concentration of industrial R&D
TABLE 3. Manufacturing R&D, 2007

<table>
<thead>
<tr>
<th>Industry</th>
<th>R&amp;D expenditures (won billions)</th>
<th>R&amp;D intensity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>21,339</td>
<td>2.97</td>
</tr>
<tr>
<td>Food</td>
<td>331</td>
<td>0.76</td>
</tr>
<tr>
<td>Textiles</td>
<td>146</td>
<td>0.86</td>
</tr>
<tr>
<td>Pulp, paper</td>
<td>40</td>
<td>0.71</td>
</tr>
<tr>
<td>Chemicals</td>
<td>2,399</td>
<td>1.49</td>
</tr>
<tr>
<td>Nonmetal minerals</td>
<td>142</td>
<td>1.20</td>
</tr>
<tr>
<td>Basic metals</td>
<td>171</td>
<td>0.63</td>
</tr>
<tr>
<td>Fabricated metals</td>
<td>75</td>
<td>1.92</td>
</tr>
<tr>
<td>Machinery</td>
<td>1,617</td>
<td>3.56</td>
</tr>
<tr>
<td>Office equipment</td>
<td>343</td>
<td>4.29</td>
</tr>
<tr>
<td>Electronic components</td>
<td>7,624</td>
<td>6.33</td>
</tr>
<tr>
<td>Communications equipment</td>
<td>2,886</td>
<td>6.71</td>
</tr>
<tr>
<td>Medical and precision equipment</td>
<td>205</td>
<td>7.50</td>
</tr>
<tr>
<td>Automobiles</td>
<td>3,831</td>
<td>3.42</td>
</tr>
<tr>
<td>Others</td>
<td>56</td>
<td>1.39</td>
</tr>
</tbody>
</table>


TABLE 4. R&D Concentration in Manufacturing, by Sector and Size of Firm, 2007

<table>
<thead>
<tr>
<th>Sector</th>
<th>Top 5 firms</th>
<th>Top 10 firms</th>
<th>Top 20 firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing</td>
<td>44.3</td>
<td>50.2</td>
<td>55.7</td>
</tr>
<tr>
<td>Chemicals</td>
<td>21.9</td>
<td>32.4</td>
<td>45.9</td>
</tr>
<tr>
<td>Electronic components</td>
<td>85.7</td>
<td>89.5</td>
<td>91.4</td>
</tr>
<tr>
<td>Communications equipment</td>
<td>64.9</td>
<td>69.4</td>
<td>73.1</td>
</tr>
<tr>
<td>Automobiles</td>
<td>67.6</td>
<td>76.4</td>
<td>88.3</td>
</tr>
</tbody>
</table>


reflects the industrial structure of Korea, which is oriented toward large firms. The top five companies in R&D investments are Samsung Electronics, LG Electronics, Hyundai Motors, Hynix, and GM Daewoo Auto and Technology, all of which are “chaebol” companies.

Another unique feature is that foreign funds and institutions have a limited role in R&D and innovation in Korea. Probably due to the once restrictive FDI policy, foreign funds seldom flow into Korea for industrial R&D. The proportion of foreign funds in Korea’s industrial R&D expenditures remains at 0.3 percent, while in many of the OECD countries, foreign funds account for more than 10 percent of domestic outlays for industrial R&D (OECD 2008a, 2008b). Given Korea’s position as one of the world’s major trading countries, it is surprising that the role of international funds and institutions in research and technology remains so low.
The Impact of the 1997 Financial Crisis on R&D and Innovation

R&D in Korea had been growing rapidly and continuously until Korea was hit by the financial crisis in 1997. R&D was severely damaged by the crisis. Industrial R&D expenditures decreased 10 percent in nominal terms from W 884.4 billion in 1997 to W 797.2 billion in 1998; R&D personnel declined 15 percent from 102,000 in 1997 to 87,000 the following year. This was a serious blow to the Korean innovation system. If the crisis had continued several more years, the Korean innovation system would have collapsed. Fortunately, Korea recovered from the crisis relatively quickly: it took only two years for industrial R&D to recover and surpass the level prior to the financial crisis (see figure 4). Two factors explain this development: one is that the government made up for the decrease in industrial R&D expenditures by increasing government R&D expenditures. The share of government in gross R&D expenditures increased from less than 20 percent before the crisis to 27 percent during the crisis. Government R&D funds flew into private industrial sectors—in particular, small technology-based firms—and helped them to maintain and expand their innovation activities. The other is the promotion of information technology (IT) and IT-related ventures, which led to an IT boom in the early 2000s. The government’s commitment to IT development is evident: the share of information technology in government R&D expenditures rose from 13 percent in 1997 to 34 percent in 2002. Such a pro-IT policy fueled innovation in the IT sector, which then spurred innovation activities in other sectors. This policy not only helped the Korean innovation system to recover its former vitality but also promoted Korea’s transition to an information society.

The financial crisis brought about two important changes in R&D and innovation in Korea. As firms—in particular large enterprises—downsized their in-house R&D activities, many of the research scientists and engineers who were let go established small-scale, technology-based start-up firms. The number of industrial R&D centers increased despite the reduction in R&D expenditures in the private sector. As a result, the role of SMEs in R&D and innovation became more important than it had been before the crisis.

Second, before the financial crisis, inward FDI had been insignificant for many years and played a minor role in the Korean economy (see figure 4). This was especially true in technology and innovation. In the early stage of development, companies acquired technology largely through informal channels; in the later stage, technology transfer in the private sector was in the form of licensing contracts rather than FDI (OECD 1996).

However, the situation began to change as Korea shifted its policy in the face of the crisis. FDI inflows into Korea increased sharply, as a result of depreciation of the local currency and asset values, the government’s deregulation and liberalization of FDI, and the investment opportunities created by corporate restructuring as well as privatization of government-owned companies. FDI companies clearly have played an important role in technology development, as their share in the Korean economy has increased.
Factors behind the Rapid Growth in R&D and Innovation

Since it started the drive to develop technology in the early 1980s, Korea has emerged as a major investor in R&D. What made this possible, and, in particular, what motivated private industries to engage so actively in R&D and innovation? Many factors have contributed to the rapid increase in private sector R&D investment, but two are paramount: demand factors and supply factors.

On the demand side, the government’s economic development projects generated huge demand for technologies. In addition, international market competition placed tremendous pressure on Korean industries to be technologically competitive. On the supply side, Korean industries were able to meet the increasing demand for R&D and innovation because they were financially able to do so and because their investments were backed by well-trained human resources. The government contributed to this development in several indirect ways, too.

Economic Development Plans Based on an Outward-Looking Development Strategy

In the early stage of development, the government’s economic development projects were the major force driving R&D and innovation in Korea. The Five-Year Economic Development Plans implemented during the period of 1961–91 specified strategic targets for industrial development. To attain the targets, companies had to invest heavily in technology acquisition, which included R&D.

Another important factor is the outward-looking development strategy adopted as a means to overcome the constraints on development, such as lack of capital and...
technology, limited market, and so on. The government’s outward-looking development strategy drove domestic industries into the international market, putting them under fierce pressure to compete with foreign companies. In order to survive, companies had to invest heavily in R&D.

Many studies have shown that the more oriented a company is toward international trade, the more it invests in R&D and innovation. A recent study at STEPI has confirmed this relationship using the Korean Innovation Survey data (STEPI 2005).\textsuperscript{10} The empirical study shows that companies with higher export intensity (export volume to total sales) tend to invest more in R&D and innovation (Shin and others 2006b). The study also finds a negative relationship between innovation activities and market concentration. In other words, companies operating in a more competitive market invest more in R&D and innovation. International competition motivates companies to invest in innovation. But a reverse relationship may also hold, as more innovative companies are more likely to compete well internationally and sell more in the international markets. In that sense, the two are mutually reinforcing, and so the effects of one on the other are determined simultaneously.

**Role of Chaebols**

On the supply side, the government’s industrial policy favoring large firms gave birth to a unique business organization in Korea: the “chaebol,” a conglomerate of businesses that is similar to the prewar “zaibatsu” of Japan. Chaebols, usually controlled by the founding families under a highly centralized structure of ownership, enjoy great financial affluence owing to economies of both scale and scope. But do chaebol companies support innovation? There are two views. One argues that in such a business structure, the major shareholder may pursue his or her private interests at the cost of those of other shareholders, is likely to seek short-term personal benefits rather than long-term company benefits, and may not actively pursue innovation. Another argues that business conglomerates like chaebol may be able to reduce transaction costs and share risks through internal transactions, while, at the same time, pooling financial resources for major investment projects. Furthermore, under the chaebol system, decision making is highly centralized, enabling the company to respond quickly to opportunities.

Only a few studies have directly investigated the differences in innovation behavior between companies belonging to chaebols and independent companies. Recently, a group of researchers at STEPI analyzed this issue using data from the Korea Information Service covering 51,270 observations for the period of 1987–2003 (Shin and others 2006a). Of the total observations, 2,064 are for chaebol companies. The study divided the period into two: before the financial crisis (1987–97) and after the crisis (2000–03), because the government changed the regulations governing the chaebol system during the financial crisis (1997–99). The new regulation bans cross-financing and cross-investment between and among chaebol companies, making the pooling of financial resources and sharing of financial risks between and among chaebol companies impossible. In other words, chaebols in the original sense disappeared toward the end of the 1990s.
The statistical analysis for the pre-crisis period shows that chaebol companies were more able and more likely to invest in R&D than independent companies. But the analysis for the postcrisis period could not find any statistically significant differences in innovation behavior between chaebol companies and independent companies. This supports the argument that chaebol companies, in the original sense, are big international operators, have deeper pockets, are able to engage in risky and expensive R&D projects that are unthinkable for independent companies, and therefore, invest more in R&D and innovation than independent companies. This finding and the high concentration of R&D expenditures in a limited number of large enterprises suggest that chaebols contributed significantly to the growth of R&D and innovation in Korea (see table 5).

**Human Resources**

Another supply factor relates to human resources. Korean industries have been able to increase R&D investments rapidly, thanks to the abundant pool of well-educated human resources. In both developed and developing countries, R&D investment is constrained more by a lack of human resources than by a lack of financing. Korea prepared itself well for development by investing heavily in education and human resource development. Reflecting the investment in education, the school enrollment rate at the tertiary level in Korea increased from 16 percent in 1980 to 37.7 percent in 1990 and to 52.5 percent in 2000. The number of full-time researchers also grew rapidly from 39,000 in 1985 to 100,000 in 1995 and to 180,000 in 2005. The number of full-time researchers per 1,000 economically active persons in Korea was 9.2 in 2007, which is lower than in Japan (10.7), but higher than in other advanced countries, such as the United Kingdom (6.1 in 2006), France (7.7 in 2006), and Germany (6.7 in 2006; see OECD 2008a).

**Government-Support Programs**

In order to promote private innovation, the Korean government offers various forms of incentives for industrial R&D and innovation. Overall, Korea employs grants

| TABLE 5. R&D Expenditures and Number of Researchers, by Size of Firm, 1997–2003 |
|---------------------------------|---------------------------------|---------------------------------|
| R&D activity                  | Small and medium enterprises | Large enterprises               |
| R&D expenditure               |       |       |       |       |       |       |
| Amount (won billions)         | 1,090.2 | 2,106.4 | 3,425.4 | 7,755.1 | 8,148.2 | 11,084.2 |
| Percentage of sales           | 2.82  | 3.14  | 3.57  | 2.07  | 1.81  | 2.05  |
| Number of researchers         |       |       |       |       |       |       |
| Total                         | 17,703 | 36,494 | 52,332 | 56,990 | 57,839 | 71,698 |
| At the doctoral level         | 474   | 1,543 | 2,291 | 3,613 | 3,878 | 5,562 |

Source: MOST various years.
and tax credits as the major instruments to promote industrial R&D; these are complemented by support programs such as procurement, technical consultancy, information, technology transfer, and so on. According to a recent survey, 259 small and large programs support private R&D and innovation (Shin and others 2006a).

How effective have these programs been in promoting industrial R&D? There have been a few attempts to evaluate the effectiveness of the policy programs (for example, Lee and Jang 1998; Shin and others 2006b). Quantitative assessments are especially needed, but full-scale quantitative analysis is not possible due to the lack of data. In 2006 STEPI undertook a survey of private industries to assess the effectiveness of the government incentive programs. The results of the survey and data from the 2005 Korean Innovation Survey formed the basis of STEPI’s econometric analysis. Both the survey results and the econometric analysis find that all categories of support programs have been effective in promoting private R&D and innovation, except the procurement program. The econometric analysis finds the tax incentive to be the most effective, followed by the loan program, human resource program, and technical consultancy. This is consistent with the OECD evaluation: “Tax incentives for R&D in Korea are generous and cover every stage: facility investment, R&D outlays, technology transfer” (Baek and Jones 2005).

The STEPI survey also identifies gaps between what private industries want and what the support programs offer. According to the survey, industries demand more government support at earlier stages of innovation. Specifically, industries need more government support at the stages of information gathering and planning (24 percent) and R&D (41 percent), while the current programs are focused on R&D (58 percent), commercialization (13 percent), and marketing (21 percent). In short, the current programs, which place greater emphasis on the later stages of innovation (commercialization and marketing), do not reflect the reality that industries need more assistance at the earlier stage (information gathering and planning).

What Korea Reaped from the Investments

Evaluations of the performance of R&D and innovation activities in Korea are mixed. Korea excels over other countries in R&D inputs, such as human resources and financing, but lags far behind in output. Moreover, R&D results have not been linked effectively to industrial uses.

Despite such criticisms, one cannot deny the positive contributions that the R&D and innovation efforts have made. Rapid growth in R&D investment has led to a remarkable increase in patent registration. The number of patents registered with the Korea Industrial Property Office increased from 1,808 in 1981 to 123,705 in 2007, an increase of almost 70 times in 26 years. Furthermore, Korea ranked fourth in the world in the number of 2007 Patent Cooperation Treaty applications, triadic patents registered in 2006, and U.S. patents registered in 2007. In the production of industrial property, Korea trails only the United States, Japan, and Germany. Of these, U.S. patents are sometimes used as an indicator of a nation’s international technological competitiveness. Only five U.S. patents were granted to Koreans in
1969, compared with 1,161 in 1995 and 6,295 in 2007, putting Korea in fourth place in the world (see table 6). According to the U.S. Department of Commerce in the late 1990s, Korea was prominent in technology areas such as information technology, pharmaceuticals, advanced materials, and automotives (Albert 1998).

Another important development is the remarkable increase in the number of scientific publications in internationally recognized academic journals. According to the Science Citation Index, the number of scientific publications written by Koreans increased from a mere 171 in 1980 to 25,494 in 2007. Korea is now the twelfth largest producer of scientific publications in the world (see table 7).

R&D efforts have also contributed to the development of high-tech industries in Korea. Korea’s technological competitiveness in semiconductors, displays, cellular phones, computers, telecommunications equipment, and so on is partly the result of the government-industry collaborative R&D.\textsuperscript{14}

Even though Korea acquired technological competitiveness in many high-tech products, its reliance on foreign core technology continues. For example, Korea succeeded in commercializing the CDMA (code division multiple access) technology, but the Korean cellular phone manufacturers paid a cumulative royalty of more than W 5 trillion to Qualcomm from 1995 to 2008 (NSTC 2009). Korea's overseas royalty payments are concentrated in the areas where Korea is known to have international competitiveness. Korea’s industrial R&D and innovation have been focused too much on commercializing foreign technologies and too little on developing original technologies, such as new materials, components, devices, and designs.

\begin{table}[h]
\centering
\caption{Number of KIPO Patents Granted and U.S. Patents Granted to Koreans, 1985–2007} 
\begin{tabular}{|l|c|c|c|c|c|c|}
\hline
\hline
Korea Industrial Property Office (KIPO) patents issued & 2,687 & 7,620 & 12,512 & 34,956 & 73,512 & 123,705 \\
U.S. patents granted to Koreans & 41 & 225 & 1,161 & 3,314 & 4,352 & 6,295 \\
Korea’s world ranking in U.S. patents issued & 24 & 9 & 8 & 8 & 7 & 4 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{14} Source: KOITA and U.S. Patent and Trademark Office.

Note: In 1981, 1,808 KIPO patents were granted.

\begin{table}[h]
\centering
\caption{Number of Korean Publications in the Science Citation Index, 1980–2007} 
\begin{tabular}{|l|c|c|c|c|c|}
\hline
\hline
Number of Korean publications & 171 & 1,227 & 2,997 & 9,124 & 12,245 & 25,494 \\
Rank in the world & — & 37 & 27 & 17 & 16 & 12 \\
\hline
\end{tabular}
\end{table}

\textsuperscript{14} Source: Korea Industrial Technology Association.

Note: — = not available.
The Link between Innovation and Industrial Competitiveness

In a neo-Schumpeterian sense, success in forging industrial competitiveness through innovation may imply a sustainable increase in the share in the market or in world trade (Cantwell 2003). Based on this, it is possible to assess the contribution of R&D and innovation to industrial competitiveness by examining how R&D and innovation affect changes in the commodity structure of exports and changes in the share of major export commodities in world trade.

As shown in table 8, in the 1980s Korea was engaged mainly in the trade of medium-low-tech and low-tech commodities and moved gradually toward high-technology commodities. Now Korea’s exports are highly concentrated in high-tech products, such as semiconductors, telecommunications equipment, displays, and so on. The share of high-tech and high-medium-tech products in Korea’s exports increased from 14 percent in the 1980s to 43 percent in the 2000s.

The contribution of individual industries to exports changed drastically from 1990 to 2007 (see table 9). The share of primary and light industries in total exports declined dramatically. The share of primary industry declined from 4.9 to 1.5 percent, and the share of textiles declined from 26.7 to 3.7 percent. In contrast, the share of high-tech products rose significantly. In particular, precision machinery, telecommunications equipment, displays, and automobiles account for almost 40 percent of total exports. The highly concentrated export structure of Korea reflects the highly concentrated distribution of Korea’s R&D expenditures and patents. Of the Korean patents registered in 2006 (KOITA 2008), electronics and communications accounted for 54 percent and machinery accounted for 15 percent. This is consistent with industrial R&D expenditures, which are concentrated in a few industries such as telecommunications, transportation, and so on. R&D-intensive industries clearly have gained market share, while low-R&D industries have lost market share. This is confirmed by

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Apparel</td>
<td>Apparel</td>
<td>Semiconductors</td>
<td>Automobiles</td>
</tr>
<tr>
<td>2</td>
<td>Iron and steel</td>
<td>Semiconductors</td>
<td>Automobiles</td>
<td>Semiconductors</td>
</tr>
<tr>
<td>3</td>
<td>Ships</td>
<td>Shoes</td>
<td>Ships</td>
<td>Telecommunications equipment</td>
</tr>
<tr>
<td>4</td>
<td>Synthetic fiber</td>
<td>Ships</td>
<td>Cell phones</td>
<td>Ships</td>
</tr>
<tr>
<td>5</td>
<td>Audio</td>
<td>Video equipment</td>
<td>Synthetic fiber</td>
<td>Petroleum products</td>
</tr>
<tr>
<td>6</td>
<td>Tire</td>
<td>Iron and steel</td>
<td>Auto parts</td>
<td>Displays</td>
</tr>
<tr>
<td>7</td>
<td>Wooden products</td>
<td>Synthetic fiber</td>
<td>Display</td>
<td>Auto parts</td>
</tr>
<tr>
<td>8</td>
<td>Miscellaneous goods</td>
<td>Computers</td>
<td>Telecommunications equipment</td>
<td>Computer</td>
</tr>
<tr>
<td>9</td>
<td>Semiconductors</td>
<td>Audio equipment</td>
<td>Computers</td>
<td>Visual instruments</td>
</tr>
<tr>
<td>10</td>
<td>Video</td>
<td>Automobiles</td>
<td>Color televisions</td>
<td>Electronic parts</td>
</tr>
</tbody>
</table>

Source: Korea International Trade Association.
World Trade Organization data, which show that Korea’s world market share increased in technology-intensive products, such as office machines, telecommunications equipment, automotive parts, and chemicals. In those areas, R&D investments also increased significantly.16

### Contribution to Economic Growth

It is hard to estimate how much investments in R&D and innovation have contributed to economic growth, because doing so involves complex data and methodological issues. One widely used method of measuring this was first suggested by Griliches and Lichtenberg (1984), who calculated the growth rate of total factor productivity (TFP) and related this to changes in the stock of R&D. Their estimates of R&D elasticity of TFP ranged between 0.17 and 0.34. Coe and Helpman (1995) conducted the same experiments using data for 22 OECD countries for the period of 1971–90. They found that the elasticity was 0.234 for G-7 countries, but only 0.07 for OECD countries, which suggests that R&D investments of G-7 countries were a lot more efficient than those of OECD countries.

Similar attempts have been made to measure the R&D elasticity of TFP in Korea. Kim (2004) estimates that R&D elasticity of TFP in Korea for the period of 1970–2002 was 0.13, while an earlier study by Shin (1996) finds a higher number, 0.166. More recently, Shin (2006) finds that Japan’s R&D is the most efficient of OECD countries.
(0.278–0.292), followed by the United States (0.167–0.263) and Korea (0.133–0.199). Korea’s R&D elasticity of TFP is almost the same as the average elasticity for all OECD countries (0.188). This suggests that R&D in Korea has performed relatively well in enhancing the productivity of the economy.

A very interesting study by Lee and others (2007) estimates the R&D elasticity in Korea separately for high-tech, high-medium-tech, medium-low-tech, and low-tech industry. They find that R&D elasticity of TFP is 0.32 for high-tech industry and 0.10 for high-medium-tech industry. However, in medium-low-tech and low-tech industries, TFP does not respond to changes in the stock of R&D—that is, R&D does not affect productivity in those industries.

Many economists use the aggregate production function to estimate the contribution of R&D stock to economic growth. The results of estimations vary depending on the periods covered, data, and methods. Lee (2008) estimates that the stock of R&D accounted for 22.7 percent of economic growth in Korea during the period of 1971–90 (see table 10). In addition, the contribution of R&D stock to economic growth has been increasing along with economic development. The share of R&D stock in economic growth during the period 1991–2006 is estimated to be 30 percent. But, sometimes, study results vary too widely to reconcile. For example, according to Hah (2004), R&D stock explains 48 percent of the growth of TFP, while Bae and others (2006) give a figure of 82 percent.

Achievements and Challenges

Korea has made enormous strides in science and technology over the past four decades. By making continuous and massive investments in R&D and innovation, Korea has succeeded in building a unique innovation system that supports sustainable growth of the Korean economy. The factors that have influenced the Korean
innovation system the most are (1) outward-looking development strategy, (2) large firm–oriented industrial policy, and (3) human resources, among many others. These are the sources of both the strength and the weakness of the system.

The strength of the Korean innovation system is its dynamism, which is fueled by the strong commitment of the government to “technology-based national development” and private industries’ efforts to be competitive. Despite the short history of R&D, Korea has harvested a rich crop of patents, scientific papers, and exports of technology-intensive products, such as semiconductors, cellular phones, liquid crystal displays, automobiles, and others.

Yet there are problems, too. Even though Korea spends a larger share of GDP on R&D than other countries, it still lags far behind advanced industrial countries in the cumulative R&D stock, which is really the determinant of a nation’s knowledge power. The challenge is to overcome the disadvantage of being a late starter.

Second, the weakness in basic sciences poses a fundamental problem for the Korean innovation system, because scientific capability determines the technological potential of a nation. Since the Korean R&D efforts have been devoted mainly to technology development, scientific research has been more or less neglected. The lack of a strong scientific base limits technological progress in Korea. The weakness in science results not just from a funding policy that favors technology development but also from weak university research capability. Therefore, strengthening university research poses a major policy challenge.

Third, excessive reliance on private industries for R&D investments has made the innovation system vulnerable in two ways. On the one hand, the system places so much emphasis on applied research and development that it has failed to build up a strong foundation for the long-term development of science and technology. On the other hand, the R&D system is too sensitive to changes in the economic and business environment. For instance, large Korean enterprises responded to the financial crisis of 1997 by cutting their R&D spending about 14 percent, almost destabilizing the R&D system. If the crisis had continued for several more years, the whole system would have collapsed.

Fourth, the Korean innovation system needs to be made more open to the outer world. Korea accounts for about 3 percent of the world R&D activities, which means that 97 percent of the world’s R&D activities are taking place outside Korea. In order to access the knowledge, ideas, and technology generated and produced outside the country, Korea needs to open the system more and promote interactions with foreign scientists and institutions. Korea’s level of international interaction is the lowest among OECD countries; if not checked, this will constitute a barrier to further growth in the future.

Fifth, the extremely high concentration of R&D activities poses a serious problem. High concentration means that only a few large firms are actively involved in R&D, while others are not. If this persists, Korean industries will be distinguished as either technologically advanced or technologically backward firms and sectors, which will make interfirm and interindustry interactions—the key elements of innovation—unlikely. This is particularly important because even large enterprises cannot sustain competitiveness without technologically strong domestic SMEs.
Lessons

Korean experiences offer some lessons for policy makers responsible for the development of education, trade, and technology in developing countries. There is no doubt that education determines a nation’s ability to absorb new knowledge and technology. Education provides individuals with initial tacit knowledge, which is an essential building block of technological learning. The government should assume full responsibility for promoting the development of human resources. Investing in education in advance, as Korea did in the 1960s and 1970s, is essential to laying a foundation for industrial development. As an economy develops, technological competence becomes critical. To build up technological competence, high-caliber scientists and engineers are needed who are capable of dealing with developments at scientific and technological frontiers. In other words, advanced education in science and technology should come first when preparing to enter a knowledge economy. In the case of Korea, education and industrialization helped to sustain and accelerate their mutual development. Education made technological learning, and therefore industrialization, possible, while industrialization enhanced the rate of return on investment in education, further promoting demand for education.

Korea’s industrialization evolved from imitation to innovation. In the initial stage, Korean industries attained technological capability through informal channels for technology transfer, such as OEM production arrangements, reverse engineering of imported machines, technical training as part of the importation of turnkey plants, and so on. Contrary to the experiences in other developing countries, in Korea FDI played a modest role in technological learning. To lay the initial technological foundation, many Korean industries resorted to nonmarket processes, relying on the ability of workers to absorb acquired technology. This approach enabled them to acquire technology at lower cost and maintain independence in their business operations. But Korea had to pay a high cost for this. It had to forgo many of the technological opportunities that foreign direct investors might have offered.

By adopting an outward-looking development strategy, the government drove Korean industries into the competitive international market, putting them under great pressure to acquire learning. Korean industries responded to these pressures by investing heavily in technology development. By developing technological competence, they have been able to survive international competition and establish world prominence in high-technology areas such as telecommunications, semiconductor memory chips, liquid crystal displays, automobiles, shipbuilding, and so on. Protectionist policy may be effective in creating initial market opportunities for domestic industries, but if such a policy is prolonged, industries will develop immunity to market pressures for innovation. It may be for this reason that export-oriented firms achieved technological learning more rapidly than import-substituting firms.

Since the early 1960s, the government has played a key role in Korea’s development. The government first initiated S&T development as part of the national economic development plan and subsequently led its development, not just as a rule setter but also as a target setter and financier. But as industrial development proceeds, it has become increasingly difficult for the government to intervene in economic as
well as R&D activities because of the increased scale and complexity of industrial activities. Therefore, the pattern of government intervention in science and technology had to change from direct involvement to indirect involvement as a facilitator and promoter. The Korean government’s policy supports have been very effective in promoting R&D and innovation in the private sector, but they also have created a culture that tends to make private enterprises overly dependent on government policy in making business decisions.

In sum, Korea owes much to its human resources and outward-looking development strategy for technological development and industrialization. The Korean experience offers two major lessons. First, human resources are the key to S&T development and thus to economic growth, and, second, nothing can motivate private businesses to invest in technology development better than market competition. For Korea to sustain the past development into the future, it has to strengthen capacity in basic scientific research and improve the framework conditions for innovation, the core of which is a competitive market.

Notes

1. Korea is a very small, resource-poor country: its land area is only 220,000 square kilometers or 99,000 square kilometers excluding the northern part, which is currently under a different political and economic system. Korea’s land, of which 75 percent is nonarable mountains, produces not a single drop of oil and contains virtually no valuable natural resources. Still it has to support a population of 70 million people.

2. This number is an estimate based on the figure for 1969, which was 5,337 (MOST 1984).

3. However, Confucianism may have adverse effects on the development of science and technology in Korea: it stresses patriotism and demands loyalty to the traditional values and, therefore, tends to devalue new, unconventional ideas. Korean society, like other Confucian societies, is less open to different ideas and systems, which works as a barrier to innovation.

4. Foreign investors also did not view Korea as an attractive place for investment. Even though Korea took a very open and liberal policy on foreign direct investment in the 1960s, few investments were made primarily because of the questions about Korea’s political stability and economic outlook.

5. For more discussions on the roles of technology suppliers and recipients in different modes of technology transfer, see Kim (1997, 100–03).

6. Another negative effect is that large-scale loans that had been brought in instead of FDI might have contributed to the financial crisis of 1997 (Chung and Suh 2006).

7. Some—for example, Amsden (1989)—say that the Korean government played the role of the market in allocating development resources during the early period of development when the Korean market system was not mature enough to function efficiently.


9. In addition to the restructuring by large firms, the government’s drive to create venture companies changed the capital market conditions for start-up companies.
10. The survey covers 2,737 companies of 23 industries for the period from 2002 to 2004. It follows the OECD manual and has been authorized as official national statistics by the National Statistics Administration of Korea.

11. There are cases where R&D investments are constrained by the shortage of suitable manpower. OECD (2003) emphasizes the importance of the supply of skilled scientists and engineers as one of the framework conditions for achieving R&D spending targets.

12. The major legal bases for the incentives are the Technology Development Promotion Law (1967) and the Industrial Development Promotion Law.

13. Despite the overall effectiveness of the programs, the STEPI study concludes that the numerous programs are not well understood by the potential users and need to be made easier and simpler to understand and access (Shin and others 2006a).

14. NSTC (2009) provides an analysis of how government R&D programs contributed to the development of major export items.

15. This follows the OECD categorization of high-tech, high-medium-tech, medium-low-tech, and low-tech.

16. In the case of office machinery, the ratio of R&D expenditures to sales increased from 5 to 6 percent in Korea during 1990–2007, contributing to the increase in Korea’s share of world trade from 2 percent in 1980 to 6 percent in 2007. More contrasting is the case of telecommunications equipment, where R&D intensity increased from 3 to 6 percent over the same period, which contributed to the change in its share of world trade from 0 in 1980 to 7 percent in 2007.

References


KOITA (Korea Industrial Technology Association). Various years. *Major Indicators of Industrial Technology.* Seoul: KOITA.


MOST (Ministry of Science and Technology). 1984. *Introduction to Science and Technology Republic of Korea.* Seoul: MOST.


NSTC (National Science and Technology Council). 2009. *Gukga R&D sung-gwa bunsuk mit sisajum* [Analysis of the Results of the National R&D Programs and Policy Implications]. Seoul: NSTC.


Roughly half of cross-country differences in per capita income and growth are driven by differences in total factor productivity associated with technological progress. What roles does innovation play in development? What makes it possible for countries like the Republic of Korea, in less than two generations, to move from a traditional agricultural society to a technologically advanced industrial society competing with advanced industrial economies? What are the factors that allow sustained technological change and productivity increases so that industrialization and development can take place over long periods? Fundamentals such as macroeconomic stability, a developed financial sector, protection of property rights, and adequate provision of public goods (such as education) are important both for investment and for innovation, but political factors such as strategies to avoid entry barriers erected by the West also play a major role (see Amsden 2001).

What are the main sources of innovation? Investment climate surveys for East Asia (Cambodia, Indonesia, Malaysia, Thailand, and the Philippines) show that more than 40 percent of enterprises cite technology embodied in new machinery or equipment (most of which probably is imported) as their most important source of technological innovation. The next two most cited sources of innovation—innovations developed in cooperation with client firms and the hiring of key personnel—were cited by 13 percent of firms, while innovations developed or adapted within the firm were cited by 12 percent of firms on average, which is only about one-quarter of the number citing new machinery and equipment (Gill and Kharas 2006). Ayyagari, Demirgüç-Kunt, and Maksimovic (2007) study the correlates of firm innovation and dynamism in a worldwide sample of firms and find that core innovation increases with firm size and with high-capacity utilization, taken to indicate high growth opportunities, while it declines with firm age (that is, younger firms are more inno-
vative). These broader measures of innovation are not closely related to income per capita, suggesting that given favorable economic and institutional conditions, firms can be highly innovative and dynamic in even the poorest economies. Formal research and development (R&D) and sophisticated innovations that lead to patents are quite different in this respect, tending to rise sharply with income per capita. There is a strong negative association between state ownership and innovation, but there is no discernible difference between whether a firm is a domestic or a foreign privately owned firm. There is also a strong association between innovation and external financing (equity finance, local or foreign-owned commercial banks, lease finance, investment funds, trade credits, and funds from family and friends), corroborating the importance of financial sector development for innovation, which is mentioned in the paper by Philip Lane. Finally, there is a positive statistical association between innovation and the extent of competition faced by enterprises.2

The Korean Experience

Korea is one of the most successful cases of industrialization and growth. Over the past 50 years, Korea has achieved remarkable growth of gross domestic product (GDP) and has performed impressively on industrial upgrading, moving into industries such as automobiles and semiconductors.

The paper by Sungchul Chung offers an overview of the role of innovation in the process of Korea’s development. In his masterful summary, Dr. Chung states that the key factors most critically linked to what he calls the Korean innovation system are (1) its relatively educated labor force, which provided a strong base for industrialization at the beginning of the process, (2) the export orientation of Korea and its outward-looking development strategy, (3) strong government intervention and control, and (4) the fact that industrial policy was heavily based on large firms (chaebol), thus allowing economies of scale and scope. Dr. Chung makes the point that Korea’s industrialization proceeded from imitation to innovation using informal channels in the first stage of industrialization with virtually zero foreign direct investment (FDI) and using only long-term loans to finance the import of capital goods and technology (which it then later reverse-engineered). There was a remarkable increase in R&D spending during the 1980s, which can be explained by world market competition (which put pressure on Korean industries to be competitive) and the ability of the system to respond to this continuous demand for innovation (the famous “treadmill” of capitalism) because a disciplined and well-trained labor force was available and both government and private industry were willing to pay for it since it was in their collective interest to do so. The result, as we know, has been the international prominence of Korea in key sectors such as automobiles, semiconductors, and telecommunications equipment.

Dr. Chung mentions three vulnerabilities. First, Korea spends a large share of GDP on R&D, but it lags far behind advanced industrial countries in the cumulative R&D stock. Second, while applied sciences and industrial technology development
have been given incentives, scientific research has been neglected. Korea is weaker in basic sciences and needs to strengthen university research. Third, the reliance on chaebols has made the innovation system vulnerable in two ways. There is no strong foundation for the long-term development of science and technology, and the R&D system reacts in a knee-jerk fashion to sudden changes in the international and national business environment. For instance, the system might have collapsed if the government had not intervened (by injecting R&D funds) to rescue it after the financial crisis of 1997.

**Learning from the Korean Experience**

There are two fundamental lessons that developing countries can draw from the remarkable Korean experience. First, at every point of its development, Korea more or less followed its comparative advantage. In the automotive sector, for example, early on, Korean manufacturers concentrated mostly on the assembly of imported parts, which was labor intensive and in line with their comparative advantage at the time. Similarly, in electronics, the focus was initially on household appliances, such as televisions, washing machines, and refrigerators, and then moved to memory chips, the least technologically complex segment of the information technology industry. Korea’s technological ascent has been rapid, but then so has its accumulation of physical and human capital, largely due to the conformity of Korea’s main industrial sectors to the existing comparative advantages and hence to changes in underlying comparative advantage.3

Second, growth was facilitated by an active industrial and trade policy. Korea protected certain sectors with high trade barriers and, in some cases, took an aggressive approach to upgrading into capital-intensive industries. Equally important, the Korean government had a record of managing protected sectors in ways that kept them subject to market discipline, which made it impossible to deviate far from the economy’s comparative advantage. Industries benefiting from protection and subsidization were required to prove on export markets that their competitiveness was increasing over time. In addition, the government worked hard to make sure that Korean manufacturers could access intermediate inputs at world prices, for example, through duty-exemption schemes and export-processing zones. So the government clearly recognized that comparative advantage mattered and that successful technological upgrading depended on firms being influenced by world prices on both inputs and outputs. The evidence indicates that the Korean government served as a “facilitating state.”

Importing technology from abroad and conducting domestic R&D are mutually supporting elements in national development strategy. It is a mistake to think that poor countries (such as Korea in the 1960s) can rely entirely on technology transfer from abroad, while more developed countries (such as Korea in the 1980s) can switch entirely to domestic R&D. Both are necessary at all levels of income, although the balance between the two may change. Even in poor economies, some indigenous
innovation effort improves the country’s capacity to absorb knowledge from abroad. As countries approach the global technology frontier, their expanded domestic R&D efforts draw even more intensively on the stock of advanced scientific knowledge in the world (see Stern, Dethier, and Rogers 2005).

**Spending on R&D**

Over the last decade, R&D spending grew much more in East Asia than in any other world region. But disparities in R&D spending between economies in the region became even wider. On the one hand, economies such as Korea, Singapore, and Taiwan, China, now devote more than 2 percent of GDP to R&D, among the most intensive R&D efforts in the world, with the business sector generally conducting more than two-thirds of R&D. China has also been rapidly boosting its R&D spending toward an official target of 1.5 percent of GDP. On the other hand, middle-income economies such as Indonesia, the Philippines, and Thailand spend a miniscule 0.2 percent of GDP on R&D, which is low relative to other economies at similar levels of income per capita.

Many studies document high social rates of return to R&D spending in developed Organisation for Economic Co-operation and Development (OECD) countries. Recent World Bank research suggests that social returns to R&D are even higher in developing than in developed countries. Why then are there such large disparities in R&D spending within East Asia and around the world? Part of the answer derives from the peculiar *nonexcludability* characteristic of knowledge, which makes it difficult for investors in business R&D to establish property rights over knowledge under the best of circumstances, but especially when the legal and institutional framework for protecting intellectual property rights is significantly weaker in some economies than in others. Being a type of investment, business R&D spending is also affected by cross-country differences in many standard factors affecting investment, for example, the extent of financial sector development, macroeconomic volatility, and the cost of capital, as well as by differences in the quality and availability of complementary factors of production, notably the level of education of the workforce (human capital) and related factors, such as the quality of academic R&D.

The main method of measuring the relationship between innovation and economic growth was suggested by Griliches and his collaborators. To measure the R&D elasticity of total factor productivity, they calculated the growth rate of total factor productivity (TFP) and related this to changes in the stock of R&D (see Griliches 1998).

The evidence (Bosch, Lederman, and Maloney 2005) shows that R&D rises exponentially with the level of development, measured by GDP per capita. Taiwan, China, and Korea in East Asia, Finland and Israel among the industrial countries, and even China and India (which experienced a takeoff that dramatically diverged from the median trajectory observed in the global data) are the “innovation overachievers.”

The literature that estimates the social rates of return to R&D for developing countries is extremely thin. The main study, Bosch, Lederman, and Maloney (2005),
suggests that R&D in developing countries is not necessarily insignificant relative to
the size of their economies and, more importantly, that returns are substantial and
generally above those for industrial countries. Exploring the determinants of R&D
across countries and over time, Bosch, Lederman, and Maloney find that the depth
of domestic credit markets, educational variables, the extent of protection offered to
intellectual property rights, the ability to mobilize government resources, and the
quality of complementary academic institutions influence cross-country differences in
R&D; taken together, a subset of these variables completely eliminates the apparent
effect of the level of development on R&D effort.

Patents

East Asia—and Korea in particular—has made huge progress in making innovations
that advance the global frontier of knowledge, using patenting in the United States as
an index. East Asian patenting per capita is closely related to patterns of R&D intensi-
ity, growing, on the one hand, in Korea and Taiwan, China, at a pace about four
times that in the developed world and reaching levels that are not too distant from
developed-country averages, while remaining, on the other hand, negligible in most
of the middle-income economies in Southeast Asia and practically nonexistent in low-
income economies. Patent citation analysis shows that not just the per capita quantity
but also the quality of patents in the most advanced innovators like Korea and Tai-
wan, China, is now approaching developed economy levels.

Analyzing the technical and scientific citations made by East Asian patents in
order to trace the international knowledge flows on which this high-level type of
domestic knowledge creation is based, Gill and Kharas (2006) find that East
Asian patented innovations continue to draw heavily on knowledge flows from
the United States and Japan. But citations to other “compatriot” patents from the
same East Asian economy or to other East Asian economies are rising quickly,
indicating that East Asian national and regional knowledge stocks are now pro-
viding an indigenous or regional foundation for new innovations and for cross-
border knowledge flows.

Policy Issues

Three main factors foster domestic innovation as well as knowledge absorbed from
abroad, namely (1) the overall business environment for innovation, including
macroeconomic stability, financial sector development, intellectual property rights,
and the quality of the information and communications technology infrastructure;
(2) human capital development; and (3) direct government support for innovation
activities, including government funding for public sector and university R&D, fiscal
subsidies and tax incentives for business R&D, fiscal incentives for FDI, and policies
aimed at promoting FDI-related technology transfers.
I do not have enough space to discuss each one of these factors in detail, but I will say a few words about education and government support, which Dr. Chung identifies as vulnerable.

Higher education is increasingly becoming a critical factor for innovation, but the efforts to improve higher education are not uniform across countries. The proportion of adults with higher education tends to rise more than proportionately with income. Korea has managed to increase higher education more rapidly than the average newly industrialized economy. Korea is also ahead in terms of quality of education (measured, for instance, by math and science scores). For example, in the OECD’s Program for International Student Assessment for mathematics proficiency, Korea has always been among the top five countries in a sample of 40.

The rationale for direct public interventions to foster domestic innovation or technology transfer from abroad (typically through targeted fiscal incentives or regulations) derives from the fact that they help to offset various types of market failures associated with knowledge, which makes it difficult for private firms to appropriate all the returns from their R&D investments and may result in the private sector undertaking inadequate innovation activities. Nonexcludability (or nonappropriability) is likely to be particularly significant for basic research that provides the early seeds for a variety of innovations by many firms or that helps countries to access the global pool of knowledge.

Support for Science and for University and Public Sector Research

A significant body of evidence indicates the positive effect of R&D funded or performed by universities and the public sector on both overall productivity and business R&D. The positive impact of university and public R&D stocks on productivity growth in OECD countries is larger than that of business R&D stocks. Nonbusiness R&D spending also has a large and significant impact on the growth of business R&D stocks in OECD countries. However, there are two caveats. First, the impact of public and university R&D is likely to depend on the quality of links between these sectors and the business R&D sector, which uses the results of more basic research to develop commercially valuable innovations and products. Second, a greater volume of public sector R&D can “crowd out” business R&D by pushing up wages for scientific and technical staff. The latter could be a particular concern in developing countries where such specialized skills are in scarce supply. At least for OECD countries, the overall impact of nonbusiness R&D on business R&D remains significantly positive, even after taking crowding-out effects into account (see Gill and Kharas 2006).

For developing countries, Lederman and Maloney (2003) have found that the perceived quality of research institutions such as universities and public research institutes has a significant, positive impact on overall R&D intensity in both developed and developing countries, as does the perceived quality of the interaction between these institutions and the private sector. They also have found that these two factors have a significant impact on the productivity of R&D in developed and
developing countries. There are significant disparities in the quality of scientific-academic research institutions and the quality of university-industry research collaboration, with Korea and the other newly industrialized economies scoring significantly higher than other Southeast Asian economies and China. To ensure that public research efforts yield good results, policymakers need to be concerned not only about adequate funding and good public-private links, but also whether public funding is allocated between various research areas using transparent, competitive, and merit-based procedures, according to criteria that strike an acceptable balance between short-term commercial interests and longer-term needs.

**Fiscal Subsidies and Tax Incentives for Business R&D**

There is a theoretical rationale for such fiscal measures as a means to counteract market failures related to knowledge. There are also serious informational and incentive problems in implementing such policies, and the limited amount of empirical work does not yield a consensus on their overall effectiveness.

Among the practical difficulties, two stand out. First, governments are unlikely to have any special information about which particular sectors might yield the largest knowledge spillovers from innovation and which therefore might merit fiscal incentives. In the face of these severe informational problems, government policies to “pick winners” could conceivably lead to outcomes that are worse than those of purely private decisions about R&D investment that, by definition, remain unconcerned about externalities and market failures (Pack and Saggi 2006). Reviewing research on the effectiveness of preferential industrial policies in Japan, Noland and Pack (2003) conclude that these policies tended to concentrate on declining sectors rather than on industries experiencing rapid technological change or increasing returns and had no noticeable impact on national or sectoral growth rates of total factor productivity. On reviewing research for Korea, they conclude, “The evidence does not support the notion that selective intervention had a decisive (or even necessarily a positive) impact on the Korean economy.” A “new” industrial policy should no longer aim to “pick winners” or sectors, but should instead be targeted at key activities that are likely to be underprovided or underperformed because of specific market failures, for example, through a generalized tax credit that does not discriminate across sectors or support for adaptation of foreign technologies to local conditions.

The second major difficulty is that a program of fiscal incentives for innovation can easily become a gateway for corruption and rent seeking. It is thus not clear if the social gains from a fiscal incentives program would offset all the compliance and administrative costs associated with it.

Cross-country experience with fiscal incentives for innovation has not been studied much until recently. In a review of the empirical literature, García-Quevado (2004) finds there is not yet a consensus on the effectiveness of public R&D subsidies. A number of studies find that such subsidies do have a significant positive impact on business R&D, but the impact declines after a certain point and even becomes negative, so that subsidies are substituting for private financing that would
have been used in the absence of the subsidy. R&D subsidies have a slightly negative impact on growth in business R&D stocks, evaluated at the mean for the sample of OECD countries. The evidence seems clearer on the effectiveness of R&D tax credits. Changes in R&D tax credits have a large impact on the user cost of capital for R&D, and the long-run elasticity of business R&D with respect to tax incentives may be substantial, on the order of \(-1\). While such analyses suggest that tax incentives are effective in stimulating business R&D, they do not, however, necessarily prove that they would be welfare enhancing overall. A full cost-benefit analysis would also need to account for the alternative uses to which the forgone tax revenues could have been put, the administrative costs of the R&D tax credit system, and the various distortions that the tax scheme could itself introduce.

Notes

1. The traditional definition of innovation processes includes not only new products and new technologies, but also activities that promote knowledge transfers and adapt production processes. Knowledge has distinctive economic features that create specialized preconditions for innovation activity. The partial nonexcludability of knowledge creates a need for specialized intellectual property rights regimes that allow inventors to recoup the rewards from risky innovation investments.

2. Ayyagari, Demirgüç-Kunt, and Maksimovic (2007) investigate the determinants of firm innovation in more than 19,000 firms across 47 developing economies. The authors find that more innovative firms are large exporting firms characterized by private ownership, highly educated managers with mid-level managerial experience, and access to external finance. In contrast, firms that do not innovate much are typically state-owned firms without foreign competitors. The identity of the controlling shareholder seems to be particularly important for core innovation, with those private firms whose controlling shareholder is a financial institution being the least innovative. While the use of external finance is associated with greater innovation by all private firms, it does not make state-owned firms more innovative. Financing from foreign banks is associated with higher levels of innovation compared with financing from domestic banks.


References


For two and a half days (June 22–24), the participants in the twentieth Annual Bank Conference on Development Economics engaged in a series of high-quality discussions and debates on the most pressing economic and development policy issues of our times. The World Bank’s Development Economics Group is grateful to the Republic of Korea’s Ministry of Strategy and Finance and the Korea Development Institute for their generous hosting of this annual international conference on development policy and research.

The timing and place of this conference were especially appropriate. As outlined at the start of the conference by Jeung-Hyun Yoon (Korea’s minister for strategy and finance), Justin Yifu Lin (chief economist of the World Bank), and Il SaKong (chairman and chief executive officer of Korea’s International Trade Association), we face great development challenges. In the short term, of course, policy makers worldwide and the international financial institutions have focused their attention on the financial and global economic crisis. This effort has already led to important institutional innovations, such as the rise of the G-20 as perhaps the premier economic grouping, huge debate on the strengthening of existing global institutions and perhaps on the need for new ones, and thinking on effective fiscal stimulus, including the opportunities for a “green” stimulus. We are in the midst of the crisis and need to increase our understanding of its causes, its impacts, how to exit, and how to change financial regulation to reduce the risk of recurrence.

However, we also face longer-run development challenges. Even before the crisis some 1.4 billion people were living in absolute poverty. Can countries implement industrial policy effectively and, if so, how? How can they accelerate technological upgrading needed for sustained growth? Can we better understand social capital and how it affects development and how it can be harnessed for favorable, rather than
negative, effects? How can slowly growing regions learn from the rapidly developing “miracle” countries of East Asia, like Korea? How do countries “take off”?

These issues and more, including climate change and the challenge of Korean unification, have been the focus of this conference, and, in the discussion of both the crisis and longer-run development issues, the focus has been on the role of the state, its interaction with market processes, and the institutional underpinnings of effective policy.

Since it is impossible to summarize a two-and-a-half-day conference in a few pages, we will therefore only highlight a few takeaways in this note.

Industrial Policy

When can industrial policy be successful? One view emphasizes the nature of the underlying political economy equilibrium—it is successful only when those implementing it have an incentive for it to succeed rather than to redistribute rents to supporters, as addressed in the papers by Jim Robinson (Harvard University) and Ha-Joon Chang (Cambridge University). Another view sees some forms of targeted policy as inevitable and indeed—as shown by Asian experience—essential for rapid sustained development, given market failures. With “good enough” political conditions, countries can implement such policies if they go about it in the right, pragmatic way. Past experience provides many lessons. Monitoring effectiveness is crucial, and export performance is less easily manipulated. There need to be strategies for exiting from failed policies, strategies that take into account the fact that the policies will themselves create new pressure groups seeking favored treatment.

Social Capital

The discussions in the papers by Partha Dasgupta (Cambridge University) and Masahiko Aoki (Stanford University) focused on the micro-foundations of social capital. The basis was seen as “trust,” which itself is rooted in reciprocity—repeated engagements in a multiple-period game. One essential parameter is the discount rate: if it is too high, future exchanges are discounted, leading to a tipping point—a breakdown in mutual enforcement and a disintegration of trust. The combination of a series of interlinked social and economic games can sustain equilibrium outcomes that are impossible with just one or the other alone. This is a useful advance in thinking about the micro-foundations of social capital. More research is needed in this area, and a better understanding of the roles of dense social networks—for example, among elites and the implications of ethnicity for such networks—is important when trying to understand the institutional basis of economic decision making. The subject has a “warm glow,” but certain types of social capital can also be harmful.
Globalization and Economic Success

In the 1960s Korea was at the level of income of Africa and often considered a hopeless case. However, Korea evolved a coherent and cohesive development policy, and incomes rose more than tenfold in about 40 years. Korea exploited the opportunities afforded by globalization; exports were 3 percent of gross domestic product (GDP) in 1960 but 38 percent by 1980. These opportunities are still there, even if in somewhat modified form. Africa differs from Korea in several respects, but there are many lessons, including the importance of nurturing small firms to provide a basis for growth and reducing the level of inefficiencies in the business climate.

Even if global growth may be slowed as households adjust demand in an effort to reconstitute their balance sheets after huge losses, the postcrisis world economy still provides opportunities for growth. Indeed, because capital is likely to be more costly for an extended period, there is a special premium on the effective use of labor and an opportunity for well-managed labor-abundant countries, including in Africa. But it is vital to secure free trade, and developing countries, which have much to gain, should be more vocal in advocacy, as discussed in the paper by Anne Krueger (Johns Hopkins University).

The themes of globalization and development were echoed in several discussions in this conference:

- **Innovation.** Innovation is greater in countries that are more open to equity-type capital inflows, especially if they have greater capability, and the development effect could be magnified after reaching a certain threshold of development per capita, as covered in the papers by Philip Lane (Trinity College, Dublin) and Sunchul Chung (Science and Technology Policy Institute, Seoul). The remarkable record of Korean innovation and its positive effects on the supply side should be required reading for other developing countries. The Korean strategy favored export-oriented firms, which were more exposed to global competition. These firms in turn benefited from government expenditure on research and development (R&D), initially absorbing almost all of the country’s R&D expenditures and now accounting for only 20 percent.

- **Crisis.** Some rethinking is required about countries’ growth strategies in a global context, as covered in the presentations by Simon Johnson (Massachusetts Institute of Technology), Olivier Blanchard (International Monetary Fund), Anne Krueger (Johns Hopkins University), and Jong-Wha Lee (Asian Development Bank). The challenges include increased vulnerability to massive and sudden capital outflows (some think some forms of controls on inflows may be necessary, but others disagree); the need to find a better reserve currency than the U.S. dollar, which may take a long time to emerge; and the need to rebalance the surplus of some countries and the deficit of the United States. However, as private consumption declines in the United States (as private saving increases), other countries will need to consume more. A big question confronting policy makers is, who will pick up the slack?
Exit strategy. Some very difficult balancing acts are needed for countries to exit from the crisis. A transition from the massive fiscal stimulus is needed. This will require significant (and often politically difficult) structural fiscal reforms. Also, there is a need to rethink monetary policy and inflationary goals, with some arguing that somewhat higher inflation targets are needed.

Financial Markets

Regulation of the financial markets clearly needs a major overhaul, as discussed in the paper by Stijn Claessens (International Monetary Fund) and the keynote speech by Simon Johnson (Massachusetts Institute of Technology). Markets are prone to crises and regulatory cycles. This can reflect reasons of political economy, as an extended period of calm reduces the perceived need for tough regulation, as those who will lose from a potential crisis (taxpayers) become complacent. Conversely, crisis periods are prone to induce overregulation, as those who will benefit in the future from less regulated and more entrepreneurial finance will not be heard, as demonstrated in the paper by Joshua Aizenman (University of California, Santa Cruz, and National Bureau of Economic Research). More regulatory independence is needed together with more disclosure to reduce information asymmetries. However, there are also operational reasons for procyclicality, for example, in market risk management, regulation itself (capital adequacy), as well as macrofinancial regulation; we still look at institutions individually.

Crises are not all bad, however: it was observed that countries that had experienced financial crises grew on average more than those that had not, a paradox resulting from their more entrepreneurial and less overregulated financial sectors. This somewhat comforting conclusion confronts two particular features of the current situation, however.

First, globalized finance has outstripped the patchwork of national regulation. There is no system of global financial governance, neither regulation and supervision nor lender of last resort nor global fiscal authority. Neither is there at present a clear alternative to the U.S. dollar reserve currency system, despite the tendency for such a system to be unstable in the long run. Neither is there a clear policy to deal with large systemic macroeconomic imbalances. Far deeper reforms are needed of the international system, especially to head off the most damaging types of recessions, those that, in part, are widely synchronized and accompanied by financial crises. The conference participants discussed these reforms in some detail, including enhanced liquidity support (reserves, International Monetary Fund enhancements, currency swaps by the central bank of the reserve currency) and the need to adapt global regulation to the conditions of developing countries. A better global system is needed to make it safer for developing countries to open up.

Second, nothing has been done to reverse the state capture by the financial sector that has taken place since the 1980s. It was argued that “finance on steroids” has outstripped its usefulness to the productive sectors and to consumers. Its power and
wealth, which were increased by deregulation, has driven an ideology that “finance is good” and resulted in institutions that are too big to fail. Nothing has been done to reverse this trend; on the contrary, with consolidation, the largest institutions post-crisis will be even larger, exacerbating the problem of moral hazard and powerful oligarchies, with enormous current and future claims on taxpayers. The answer, as in previous oligarchic cycles, is to break up the largest banks.

But no such fundamental steps have yet been taken, and the outlook is not promising, especially as many of the current features of the crisis, including outrage at bankers, had been features of previous crises. As illustrated in the paper by Giovanni Zanalda (Duke University), societies have the opportunities to draw the right conclusions from crises, but they often fail to do so.

Other topics were covered in a range of interesting papers and discussed at parallel sessions:

- Decoupling or recoupling of East Asia (change from unidirectional—North-South—dynamics to a bidirectional relationship)
- Relationship between exporting and productivity growth (self-selection versus “learning by doing,” which is stronger for firms with some capability)
- Implications of production networks
- Household and firm responses to shocks
- Implications of adaptation of the euro
- Climate change
- Korean unification

The discussion on climate change noted the need to “turn the discussion upside down” with a shift in focus from technical and physical issues to social ones, with a strong emphasis on people and a local stress on adaptation. It also noted the gap in trust between developed and developing countries in the global architecture—the reluctance to recognize the “polluter-pays” principle and the lack of credibility that existing shortfalls from aid commitments create for promises of additional support for climate change. The apparent gap in “trust” complicates efforts to find a globally optimal solution to the lack of adequate global social capital.

On the question of Korean unification, comparisons between Korea and Germany show that the difficulties of a similar unification process—unification before transition—would be far greater in Korea; the Koreas are more divergent than the Germanys were and more equal in size. However, they also show that such a process is very unlikely given the political conditions. Far more likely is a paced transition followed by extended unification. This would be a very different process, and possibly a win-win one, but there would still be severe stresses; for example, even with “two countries, two converging systems,” Korea’s constitution will make it impossible to block labor flows from the North as the regime is liberalized.
Conclusions

This summary can only offer a superficial picture of some of what transpired at this very high-quality conference. The exchanges that took place led to many useful thoughts and suggestions on how to face the formidable policy and development challenges facing all developed and developing countries in the world today. Finally, many of the sessions will provide inspiration for research in future years.
Vice President Lin of the World Bank, honorable Ambassador Vargö of Sweden, and distinguished guests, it is now time for us to end the stimulating discussion of the past three days and to close the 2009 ABCDE. Thanks to your active participation and the hard work by our organizers, we are ready to conclude a very successful conference. On behalf of the Korean government, I express my deep appreciation to all of you. The past three days have, indeed, been a valuable learning experience for us. The conference has offered a golden opportunity to listen to the insightful analysis and outlooks on the global economy of world-renowned scholars and prominent presenters. This conference included presentations and discussions not only on the cause of the economic crisis and ways to tackle it, but also on many future challenges such as climate change, sustainable growth, and financial regulatory reform. I believe that the expansive scope and extent of the discussion have been the greatest accomplishment of this year’s conference. I sincerely hope that the ABCDE will continue to serve as a venue for convening intellectuals from around the world to discuss the present and future of the global economy.

Distinguished guests, we are experiencing a global economic crisis unprecedented in scope. Fortunately, however, history has taught us that a crisis is always accompanied by change, and change is accompanied by new opportunities, and such opportunities will only be captured by those who have been carefully preparing for them. Through this conference, I hope that you have learned how to prepare for and seize the opportunities that will emerge when the crisis is over. In Buddhism, there is a saying, “Those who meet must part and those who part must meet again.” Although we are about to finish our conference, I am filled with anticipation and look forward to meeting all of you again at the ABCDE 2010 in Sweden. When we meet again, I am sure that the crisis will have been tackled successfully and that hope will have been restored.

Young Geol Lee is the vice minister of strategy and finance of the Republic of Korea.
Once again, I would like to thank the organizers, the World Bank, the Korean Development Institute, my staff at the Korean Ministry of Strategy and Finance, and the Intercom Convention Service for the successful conference. I hope all of you have a safe trip back to your homes. Thank you.
ECO-AUDIT

Environmental Benefits Statement

The World Bank is committed to preserving endangered forests and natural resources. Annual World Bank Conference on Development Economics 2010, Global: Lessons from East Asia and the Global Financial Crisis is printed on recycled paper made with 50-percent post-consumer waste. The Office of the Publisher follows the recommended standards for paper usage set by the Green Press Initiative, a nonprofit program supporting publishers in using fiber that is not sourced from endangered forests. For more information, visit www.greenpressinitiative.org.

Saved:
- 8 trees
- 3 million BTUs of energy
- 778 pounds of net greenhouse gases
- 3,746 gallons of waste water
- 227 pounds of solid waste
Themes and Participants for the

ANNUAL WORLD BANK CONFERENCE ON
DEVELOPMENT ECONOMICS

Stockholm, Sweden

“DEVELOPMENT CHALLENGES IN A
POST-CRISIS WORLD”

May 31–June 2, 2010

Environmental Commons and the Green Economy
Post-Crisis Debates on Development Strategy
The Political Economy of Fragile States
New Ways of Measuring Welfare
Social Programs and Transfers: Are We Learning?

Anders Borg • Gunilla Carlsson • Justin Yifu Lin • Elinor Ostrom • Thomas Sterner • Ramon López • Simon Levin • Abhijit Banerjee • Shang-jin Wei • Geoffrey Heal • Franjo Štiblar • Torsten Persson • Stephen Ndegwa • James Fearon • Alan Gelb • Louise Anten • Joseph Stiglitz • Justin Wolfers • Carol Graham • Peter Hammond • Leonardo Gasparini • Partha Dasgupta • Esther Duflo • Ingrid Woolard • Michael Clemens • Jane Fortson • Ann Harrison • Joakim Stymne

The Annual World Bank Conference on Development Economics (ABCDE) is one of the world’s best-known series of conferences for the presentation and discussion of new knowledge on development. The conference provides a forum for the world’s leading development thinkers to share new knowledge and ideas. “Lessons from East Asia and the Global Financial Crisis” was the theme of the ABCDE held in Seoul, Republic of Korea, on June 22–24, 2009. The conference was co-organized by the Government of the Republic of Korea, the Korea Development Institute (KDI), and the World Bank.

IN THIS VOLUME:

http://www.worldbank.org/abcde