Creating a Learning Society

KENNETH J. ARROW LECTURE SERIES

KENNETH J. ARROW LECTURE SERIES

Kenneth J. Arrow's work has shaped the course of economics for the past sixty years so deeply that, in a sense, every modern economist is his student. His ideas, style of research, and breadth of vision have been a model for generations of the boldest, most creative, and most innovative economists. His work has yielded such seminal theorems as general equilibrium, social choice, and endogenous growth, proving that simple ideas have profound effects. The Kenneth J. Arrow Lecture Series highlights economists, from Nobel laureates to groundbreaking younger scholars, whose work builds on Arrow's scholarship as well as his innovative spirit. The books in the series are an expansion of the lectures that are held in Arrow's honor at Columbia University.

JOSEPH E. STIGLITZ AND
BRUCE C. GREENWALD

CREATING A LEARNING SOCIETY

A New Approach to Growth, Development, and Social Progress

READER'S EDITION

Columbia University Press
New York

Columbia University Press
Publishers Since 1893
New York Chichester, West Sussex
cup.columbia.edu

Library of Congress Cataloging-in-Publication Data

Stiglitz, Joseph E.
Creating a learning society : a new approach to
growth, development, and social progress /
Joseph E. Stiglitz and Bruce C. Greenwald. — A reader's edition.
pages cm. — (Kenneth J. Arrow lecture series)
Includes bibliographical references and index.
ISBN 978-0-231-17549-4 (pbk. : alk. paper)
1. Social learning. 2. Information society. 3. Progress.
I. Greenwald, Bruce C., 1946- II. Title.

HQ783.S6942 2015
303.3'2—dc23

2015002293

∞
Columbia University Press books are printed on permanent
and durable acid-free paper.

This book is printed on paper with recycled content.
Printed in the United States of America

c 10 9 8 7 6 5 4 3 2 1

Cover design: Noah Arlow

Contents

reader's edition grant by
Figure Foundation

determination in wisdom

Preface to the Reader's Edition

IN THE few months since this book first came out, it became clear that we had hit a resonant chord. The book, growing out of a lecture in honor of one of the world's greatest economists, Kenneth J. Arrow, was aimed at a more academic audience than our other, more popular writings. We unabashedly included complicated mathematical equations; mathematics has become the language among academic economists, and we wanted to speak to them—to persuade them that some of their most long held precepts, even the virtues of free trade, at least for developing countries, needed to be rethought. We wanted to challenge them to think more deeply about what had really brought the enormous increases to standards of living that have marked the last two hundred years: the creation of learning societies. Governments, we argued, should focus on what creates a learning society. Some of the policies which economists had traditionally argued for actually impeded that.

In the last two decades it had become conventional to describe the economy toward which we were moving as a knowledge economy and an innovation economy. But little attention was given to what that meant for the organization of the economy and society—or even for narrower subjects like economic policy. But our argument here was more general: creating a learning society is necessary for advancements in standards of living, even for economies well below the *frontier* that were not at the vanguard of advances in science and technology.

Accordingly, the reception that the book received—not just in advanced countries in Europe, but in developing countries and emerging markets—was heartening. So too in places like Malaysia, Singapore, Turkey, Jordan, South Africa, and elsewhere where we had an opportunity to discuss the ideas. A major Dutch think tank, closely linked to the government, even released a report, *Towards a Learning Economy,* a blueprint for that country going forward.[1]

Many editors, readers, and scholars from both the advanced countries and the emerging markets asked if we could produce a shorter version focusing on the core theoretical developments, the main messages, and the central policy prescriptions. This edition responds to that request. In the almost one year since we finished the draft of the first edition of the book, we have continued with our research, sharpening some of the results, clarifying some of the complex trade-offs, and linking some of the ongoing policy debates to our overarching framework. This edition incorporates some of these new ideas, especially in chapter 11.

Chapters 5 and 6 in the previous edition described the relationship between competition and innovation and discussed the efficiency of the market in innovation; in this reader's edition, these chapters have been basically rewritten, but the message is the same: the relationship between competition and innovation is complex—far more complex than anyone previously realized. Still, we can identify some of the critical factors in this relationship (e.g., the role of government). Moreover, there is no presumption that the market is efficient in either the pace or direction of innovation. Insights from our analysis of the factors affecting the pace of innovation can and should help shape our innovation policies.

In the original version of the book, part 2 was devoted to developing the mathematical analytics underlying "creating a learning society." But the basic insights can be conveyed in words, and chapters 7 and 8 attempt to do just that with the help of appendices providing simple diagrammatic expositions of the basic mathematical ideas. Readers interested in the formal models underlying the analysis are referred back to the original edition of the book and to the more extended papers that we cite.

The original version included commentary on the Arrow lecture by Robert Solow, Kenneth Arrow, and Philippe Aghion; a summary of the discussion that followed the Arrow lecture; and a complementary paper on industrial policy by Aghion. Because of space limitations, these have been omitted from this edition.

In addition to the acknowledgments listed in the preface to the original edition, we wish to acknowledge the assistance of Eamon Kircher-Allen and Feiran Zhang in preparing this edition of *Creating a Learning Society*.

Joseph E. Stiglitz
Bruce C. Greenwald
New York, 2015.

Acknowledgments

Acknowledgments for the Series

The Kenneth J. Arrow Lecture Series has been made possible through the efforts of Columbia University's Committee on Global Thought (which I chaired at the time the series was inaugurated, and which is now co-chaired by Saskia Sassen) and by the Program in Economic Research (PER) of the Department of Economics at Columbia University (chaired by Michael Woodford) with the support and encouragement of the Columbia University Press.

We are especially indebted to Robin Stephenson, Gilia Smith, and Sasha de Vogel of the Committee on Global Thought and Myles Thompson and Bridget Flannery-McCoy of the Press for guiding this book to publication.

We are grateful for the support of the Kaufman Foundation, who funded the Inaugural Lecture.

Acknowledgments for the First Arrow Lecture

While this lecture represents a continuation of my extraordinarily fruitful collaboration with Bruce Greenwald, extending now for more than three decades, my work on the topics covered here goes back to my days as a graduate student, with a half-century of accumulated debts. I need to begin by especially thanking two of my teachers, who were present at the lecture at which the earlier version of these ideas was presented, Kenneth Arrow and Robert Solow. At the time I was a graduate student, growth theory was all the rage, and *endogenous* growth theory—explaining the rate of technological progress—absorbed much of our attention.

I was fortunate to be able to spend time not only at MIT but also at the University of Chicago and Cambridge University. At MIT, I should acknowledge the influence of Evsey Domar, Paul Samuelson, and Franco Modigliani; at Cambridge, of Nicholas Kaldor, James Meade, James Mirrlees, David Champernowne, and Frank Hahn. Hirofumi Uzawa brought a group of us together from throughout the United States for a summer at the University of Chicago to discuss these issues—including George Akerlof, Eytan Sheshinki, and Karl Shell. While at Cambridge, I began my collaborations with Tony Atkinson and Partha Dasgupta, and the ideas developed with them permeate this book. Other coauthors whose ideas are discussed extensively in this book and who have influenced my thinking on the topics discussed here include Karla Hoff, Raaj Sah, Barry Nalebuff, Carl Shapiro, Richard Arnott, Avi Braverman, Andrew Charlton, Mario Cimoli, Michael Cragg, Domenico Delli Gatti, Giovanni Dosi, Avinash Dixit, Shahe Emran, David Ellerman, Jean-Paul Fitoussi, Drew Fudenberg, Jason Furman, Mauro Gallegati, Richard Gilbert, Sandy Grossman, Geoff Heal, Thomas Hellman, Claude Henry, Arjun Jayadev, Kevin Murdoch, David Newbery, Akbar Noman, Jose Antonio Ocampo, Michael Salinger, Marilou Uy, Alec Levinson, and Andy Weiss.

And then there are the host of others with whom I have had discussions on the topics included here, who have helped shape my thinking, and whose influence can be felt in the pages that follow. These include Alice Amsden, Amar Bhide, Hans Binswanger, Ha-Joon Chang, Paul David, Bob Evenson, Gene Grossman, Peter Howitt, Mort Kamien, Justin Lin, Glen Loury, Edwin Mansfield, Richard Nelson, Takashi Omari, Edmund Phelps, Rob Porter, Hamid Rashid, Jerome Reichman, Dani Rodrik, Paul Romer, Nancy Schwartz, Michael Spence, Robert Wade, Michael Whinston, and Sidney Winters.

We need to extend a special thanks to the participants in the lecture, and especially the commentators, including Philippe Aghion, Kenneth Arrow, Robert Solow, and Michael Woodford.

For this volume, the authors wish to express their gratitude to Julia Cunico, Kevin Findlan, Laura Morrison, Hannah Assadi, and most importantly, to Ava Seave and Anya Schiffrin.

I should also acknowledge the team of research assistants who helped prepare the book. Special thanks go to Laurence Wilse-Samson and Eamon Kircher-Allen for their much appreciated assistance in preparing this manuscript. Sandesh Dhungana, An Li, Erin Kogan, and Feiran Zhang also provided valuable input.

Joseph E. Stiglitz

Creating a Learning Society

Preface to the Original Edition

THIS VOLUME is the result of the first in a series of lectures to honor one of Columbia University's most outstanding graduates, Kenneth J. Arrow, who received his Ph.D. from Columbia in 1951. His thesis, later published as *Individual Choice and Social Values*, was a landmark in economics, philosophy, and political science. In the more than sixty years that followed, Ken went on to become a giant in economics, political science, organization theory, and operations research.

Columbia University has had a long line of distinguished graduates and faculty—including six Nobel Prize winners in the past thirteen years. The faculty list for economics includes Milton Friedman, who taught at Columbia for ten years; Arthur Burns, who served on President Eisenhower's Council of Economic Advisers from 1953 to 1956 and as chair of the Federal Reserve Board from 1970 to 1978; and Wesley Mitchell, who along with Burns played a central role in founding the National Bureau of Economic Research, one of the nation's most important think tanks, which focused in its earlier years on enhancing understanding of economic fluctuations. There are a host of other greats, known more to those within the economics profession than those outside, including Harold Hotelling, Albert Hart, and John Bates Clark (whose namesake medal is awarded every year to the economist under forty who has made the most significant contribution to economics; Arrow was the fifth recipient of the honor).

With all of these potential luminaries, our decision to honor Kenneth Arrow was easy: No individual has done more to change how we think about economics—and about society beyond economics—during the past six decades. In a sense, virtually all theorists—and most policymakers—of our generation are students of Arrow (and, it might be added, our students can be considered their "grandstudents"). Ideas that he first put forward a half-century ago have permeated our thinking.

A lecture series like this provides the opportunity to approach issues a little bit more expansively than one is able to in journal articles. When we initiated the series, we had hoped that it would open up a lively discussion about a variety of areas within economics, political science, and philosophy. The Committee of Global Thought spans multiple disciplines, and Arrow is one of the few scholars of recent decades whose work has cut across fields, having profound implications on each. One of the reasons why it is a particular pleasure to have Ken Arrow as the honoree of this lecture series is that we hoped to focus every year on one aspect of Ken's work. Since Ken has written about so many different areas, this would make the lecture series broad, engaging people from throughout the university community.

The lecture series has lived up to our hopes. In the first lecture, in late 2008, Bruce Greenwald and I focused on one aspect of Arrow's contribution to our understanding of growth: how technological progress is related to what we do. It was, in a sense, the founding paper in what has since blossomed into a huge literature on endogenous growth, where the pace of innovation is the central object of study.

The second lecture took up Arrow's seminal thesis, in which he asked a question of greater generality than had ever been posed—and academia has struggled to come to terms with the disturbing answer that he provided. Nearly two hundred years earlier, the great French mathematician Nicolas de Condorcet had shown that a democracy, making a choice among three alternatives by a majority vote, might not be able to reach a determinate answer. Alternative A might be preferred over B by a majority, B over C by a majority, and C over A by a majority. Under a set of plausible hypotheses, Arrow showed that this problem could arise with any voting mechanism (with the obvious exception of giving all decision-making power to a single individual).

The implications of this—and the conditions in which this seeming paradox might not hold—were discussed in the Second Annual Arrow Lecture, delivered at Columbia University on December 11, 2009, by two distinguished Nobel Prize winners who have devoted considerable intellectual energies to understanding the Arrow Impossibility Theorem, Eric Maskin and Amartya Sen.

In 2010, we turned to his contributions to financial markets, with a lecture by Jose Scheinkman, then of Princeton University and now of Columbia (with discussions from Patrick Bolton of Columbia University and Sanford Grossman).

The 2011 lecture focused on Arrow's contributions to the environment, and climate change in particular, with a lecture by Sir Partha Dasgupta, and discussions by Geoffrey Heal and Scott Barrett, both of Columbia University. In 2012, Amy Finkelstein of MIT, along with discussant and MIT colleague Jonathon Gruber, continued Ken Arrow's pathbreaking work in the economics of health, a paper written forty-seven years earlier, whose influence continues today, and which was also a pathbreaking paper in the broader area of the theory of moral hazard.

In 2013, we returned once again to climate change, with a lecture by Christian Gollier of the Toulouse School of Economics entitled "Pricing the Planet's Future: The Economics of Discounting in an Uncertain World," with discussions by Bernard Salanié of Columbia, Stiglitz, and Arrow.

What made each of these occasions so exciting—and moving—was Arrow's participation and his reactions to these lectures inspired by his own work.

What also made these occasions moving was that the speakers had not only a strong intellectual bond with Arrow but also close personal ties—sometimes as students, often as colleagues. No one that we have approached to give the Arrow Lecture has ever turned us down—everyone, as busy as they are, went to great lengths to rearrange their schedules so they could have this opportunity to show their respect and honor for one of the century's great economists. Each delivered a lecture worthy of the person they were honoring.

The inaugural lecture of the series, on November 12, 2008, was an especially important event, because it brought together Ken Arrow and Robert Solow, two of the economists who were responsible for creating a new field of economics—growth theory—perhaps the most important area in the decades immediately after World War II. The events gave them an opportunity to reflect on what has happened to the subject in the half-century since their seminal contributions. Philippe Aghion of Harvard supplemented his comments on the lecture by offering his observations on industrial policy (one of the main topics of the talk) in the paper published here, "The Case for Industrial Policy."

The original lecture has been extended (partly at the suggestion of Solow and Arrow) into a fuller treatment of the subject. In the original lecture, Bruce Greenwald and I had focused our attention on showing how Arrow's insights into learning necessitated rethinking one of the most fundamental tenets of modern economics, the virtues of free trade. We showed that there was an infant-economy argument for protection. Solow and Arrow observed that our analysis showing the desirability of government intervention into the market applied equally forcefully to a closed economy, without trade. We publish here their comments on our original lecture. The research that we did subsequently, and report here, shows how right they were.

This volume begins with introductory remarks by Bruce Greenwald and me, our personal tribute to Ken, showing our affection and respect.

Joseph E. Stiglitz,
New York, 2014

Introduction

JOSEPH E. STIGLITZ AND BRUCE C. GREENWALD

IT WAS a real pleasure for us to deliver the First Annual Kenneth J. Arrow Lecture at Columbia University—to honor our teacher, someone who has had a lifelong influence on our thinking, as he has had on an entire generation of economists.

There is, in fact, a sense in which everyone in our generation was a student of Kenneth Arrow—even those who were not fortunate enough to take his class. His ideas influenced us, as did his style of research and his breadth of vision. He is a true model of a scientist. He could provide the definitive proof of the Pareto optimality of the competitive equilibrium (the first fundamental theorem of welfare economics), then go on to explain why the assumptions were wrong—and then go on to develop models incorporating more realistic assumptions, overturning the earlier conclusions about the efficiency of the market.

Both Arrow and Robert Solow, another of our teachers that our lecture honored, performed just those kinds of analytical feats in a series of papers that inspired this volume. The first was a paper that Solow wrote in 1956, which showed that an increase in the savings rate would *not* lead to an increase in the long-run growth rate—that was determined by the rate of productivity growth. Then, in 1957 he decomposed the sources of economic growth and argued that most of economic growth was related not to increases in factors of production—like labor and capital—but rather to increases in productivity. Before that, economists focused on savings and capital accumulation, but not on the role of technological progress, as the source of the enormous increases in our standard of living over the past two hundred years.

In 1962 Ken Arrow published two important papers attempting to explain technological progress. One focused on research and development (1962b) and the other on learning by doing (1962a). This latter paper observed that, in the process of producing and investing, one learns. As we produce and invest, we get better at what we do. If one builds more ships, one becomes more efficient at building ships. Productivity increases. This was one of the earliest papers on what has come to be called endogenous growth theory, where the pace of innovation is determined within the model.

Each of the Arrow lectures is intended to build off one of Arrow's pathbreaking contributions. For our lecture, we took his work on innovation, in particular his remarkable 1962 paper on learning by doing. That paper itself is in part a commentary on an earlier important Arrow paper. Two hundred and forty years ago Adam Smith talked about the efficiency of the competitive market economy. He argued that competitive equilibrium was efficient, that the pursuit of self-interest would lead, as if by an invisible hand, to the well-being of society. It took a long time for economists to determine in what sense that was true (what economists now refer to as Pareto optimality) and the circumstances under which it was true. The critical works proving the conditions under which competitive equilibrium was in fact Pareto efficient were Arrow's (1951b) and, contemporaneously, Gerard Debreu's (1952; also Arrow and Debreu 1954).

Arrow had assumed in that paper that technology was fixed, that is, that there was no innovation.[1] His paper on learning by doing challenged that assumption. For a modern economy, innovation is clearly central. In that paper, as well as in his other 1962 paper on R & D (1962b), Arrow explained why the production of knowledge is very different from the production of conventional goods.

When technology is endogenous, markets are not, in general, efficient. But this immediately raises a further question: How should government intervene in the market to enhance efficiency and societal welfare? Remarkably, in the fifty years since Arrow's 1962 paper, that question has been addressed only in a piecemeal way (e.g., in discussions about intellectual property and patent policy).

In our lecture, we investigated the implications of learning by doing for the long-standing presumption in favor of free trade. It made for a good lecture topic, which gave way to a day of useful discussions and interesting reactions, many of which are included at the end of this volume. But as we prepared our lecture for publication—since the Arrow lectures' inception, a book series was planned to accompany them—and we took to heart the comments made by Arrow and Solow, it became clear to us that to do justice to the issues we had raised required more than a short lecture. Arrow's work had opened the door to a large body of fresh analysis on how to create a learning economy and society—and how government can and should intervene to improve societal well-being.

That we chose Arrow's learning perspective as the foundation for our lecture—and the subsequent elaboration that resulted in this volume—is neither coincidence nor contrivance. Rather, Arrow's work proved the perfect starting point for the same reason that the lecture series was named after him: The contributions he made to the field are still so important that half a century later they are often the ineluctable jumping-off point for present-day work.

Like other great economists of his generation (including Solow), Arrow has ultimately been interested in improving the practice of economic policy. Clarifying economic thinking, while valuable in itself, really accrues value in the course of being applied to particular situations where policy decisions are being made, in some cases being made badly, almost always in ways that can be improved upon. In approaching the question of free trade from Arrow's learning perspective, not only do we honor his legacy and challenge the conventional views, but also hopefully we make a contribution to a key set of policy issues: how to increase the pace by which living standards increase, especially in developing countries.

The fact that markets on their own are not efficient when innovation is endogenous raised the question which is at the heart of our lecture and the book to which it gave rise: What should be the role of policy in promoting economic efficiency? Advocates of unfettered markets often respond to this question by championing the market's ability to innovate. But there is remarkably little systematic inquiry into whether markets generate the optimal level and form of innovation. Our lecture was intended to fill this gap, with specific applications to trade policy.

There was at the time we delivered our lecture already a long-standing exception to the presumption in favor of free trade in the idea that it might be appropriate to protect infant industries. Thus, if a particular industry grew with protection, and got stronger as it grew because it benefited from economies of scale, one might think there was an argument to protect that industry. There is a second exception to the principle of free trade associated with price manipulation. If a country has a large industry in the world economy, then it can manipulate the terms of trade (that is, international prices) to its benefit. These two exceptions are related, and under careful scrutiny the second argument enhances our understanding of the limits of the first: if one does not alter the terms of trade, it does not matter where the protected industry develops. Nigeria might, say, protect its auto industry until it was strong enough to compete in global markets. But if that industry can efficiently develop in England—and as long as the import prices reflect the productivity gains—Nigerians will benefit by buying and importing those cars just as much as people in England do.

In fact, the terms-of-trade argument has always been a fairly weak argument. The argument that countries, even the United States, can move the terms of trade is difficult to make in practice. Thus, the standard theories do not provide very persuasive reasons for trade interventions. Nevertheless, there seems to be a persistent pattern of *successful* economies practicing trade restrictions.

In thinking about this problem, we applied Arrow's lessons in a way that brought us to a different conclusion, which forms the heart of this lecture and book. Our analysis shows that these successes are not based on the infant-industry argument for protection, where there are benefits *within* an industry to learning by doing. Instead, there is an infant *economy* argument for trade interventions. The intuition is remarkably simple: We explain why innovation is likely to be more centered in the industrial sector rather than the agricultural or craft sector. The industrial sector is not only better at learning, but it also generates more externalities—more learning benefits—than the rest of the economy. An economy that starts out without a strong urban industrial sector—one that is importing those goods—is unlikely to develop improvements in productivity, even within that sector. There is little learning, little innovation. Trade barriers are necessary to enable that economy to develop those industrial enterprises even though it might seem inefficient to do so in the beginning, because it runs counter to the country's current comparative advantage.

So far, the argument runs parallel to that of the standard infant-industry argument. But here is where the difference arises: enterprises in protected industries will generate productivity growth not only in their sector but also across different individual products within that sector—and also across agriculture and other sectors of the economy. It is the externalities generated by the sector that provide the real rationale for intervention.

The classic example of this is, of course, the Agricultural Extension Service in the United States, where principles of industrial research got applied to farms in an extraordinarily efficient manner. As much as anything, it accounts for the remarkable growth in agricultural productivity in the United States.

That was the basic idea that we proposed in our lecture. It calls for a kind of protection that is not industry specific. The classic complaint about infant-industry arguments—that trying to pick successful industries is a doomed effort—does not apply. This is an argument for a broad set of tariff barriers (or exchange rate interventions), within which one hopes that the best industries will survive and prosper.

A Guide to This Volume

In the years following the first lecture, our ideas took on new life. As we worked the ideas of our lecture into different papers and continued our research on related topics, it became clear to us that we had more than a slim conference volume's worth of material. Our lecture on "creating a learning society" was growing into a full-fledged body of theory that required historical context, examples of general and specific applications, and discussions of political economy. With that realization, this book began to take shape. The result is something far more expansive than the original lecture, though the core intellectual inspirations for the book are the same as those that guided us in 2008.

In the first few chapters of this book, we lay out our basic theses: that most of the increases in standards of living are, as Solow suggested, a result of increases in productivity—learning how to do things better. And if it is true that productivity is the result of learning and that productivity increases (learning) are endogenous, then a focal point of policy ought to be increasing learning within the economy; that is, increasing the ability and the incentives to learn, and learning how to learn, and then closing the knowledge gaps that separate the most productive firms in the economy from the rest. Therefore, *creating a learning society should be one of the major objectives of economic policy*. If a learning society is created, a more productive economy will emerge and standards of living will increase. By contrast, we show that many of the policies focusing on static (allocative) efficiency may in fact impede learning and that alternative policies may lead to higher long-term living standards. Thus, the theory that we develop provides one of the most compelling and fully articulated critiques of the Washington consensus policies that dominated development thinking in the quarter century before the Great Recession. The theory also provides the basis of a new theory of the firm—a new answer to the question posed more than 75 years ago by Ronald Coase: What determines the boundaries of firms, what goes on inside the firm? It also provides a new approach to thinking about both static and dynamic comparative advantage.

Part 1 also gives the reader a view of the historical, empirical, and theoretical background and justification for our learning-society perspective. We describe key aspects of creating a learning society: the processes and determinants of learning and some of their broad implications for economic architecture—the design of the economic system and its subcomponents (most importantly, firms)—and policy. We explain the implications of "localization of knowledge" (both technologically and spatially), extend the concept of learning by doing to learning to learn by learning, explain why geographically concentrated large enterprises, traditionally in the industrial sector but more recently in the modern services sector, have been at the center of growth—with high rates of productivity increases and large spillovers to other sectors of the economy. We explain, too, the link between macro-stability and long-run productivity growth—a new rationale for why *real* macro-stability is so important.

Having analyzed the basic determinants of learning, we address two critical questions: Is there likely to be more or less learning in economies that are more competitive (with more firms)? And is the market likely to be efficient in the level and pattern of innovation and learning? In asking the latter question, we note that the level of competition (concentration) is itself *endogenous*—though it can be affected by government policies. As we have already noted, Arrow's earlier work provided more than a little hint that the outcomes of market processes would not be efficient, though he did not directly challenge Schumpeterian views which championed the innovative virtues of the market. The picture that emerges from our analysis is complex: Joseph Schumpeter was overly optimistic about monopolies—he thought that they would be only temporary and that competition to be the dominant firm drove innovation. We show that monopolies may be far more persistent than he (and latter-day Schumpeterians) thought and that the fight to be the dominant firm may be far less effective in stimulating innovation than he thought. Still, Schumpeter was right that more competitive markets, with many small firms, are likely to be *less* innovative.

The central message that emerges is that there is an important role for government to play in shaping an innovative economy and in promoting learning. The later chapters in the book explore in more detail how the government can best do this.

Chapters 7 and 8 provide the key analytical results, moving from simple models to more complex. Chapter 7 looks at a two-good (agriculture and manufacturing) closed economy (no trade) model and explains how policies promoting the industrial (manufacturing) sector (such as subsidies) lead to higher rates of growth and welfare. The short-run (allocative) distortions are more than offset by the long-term learning benefits. Simple formulae describing the optimal subsidy are derived. In this simple setting, we can compare the rate of innovation if there is competition with that when the industrial sector is dominated by a single firm. Innovation will be higher with monopoly, but whether welfare will be higher is ambiguous and depends on learning elasticities and discount rates.

Chapter 8 extends the analysis to an open economy, establishing the infant-economy argument for protection. Because the industrial sector not only has a greater capacity for learning but also more learning spillovers, encouraging that sector through protection or industrial policies can lead to higher growth and societal welfare. The force of the argument for protection is much weaker in developed economies. In economies like the United States, Europe, and Japan, there is already a dense infrastructure that has the scale to develop ideas and innovations, though there may still be cross-sector or cross-industry learning externalities that might warrant government intervention.

The theory has a wide range of implications. To illustrate: Our analysis suggests that it is desirable for large groups of countries to work together to facilitate trade amongst each other, while erecting certain barriers to trade from the outside. Competition and incentives matter. Having broad collections of countries, like the European Union, competing behind broad barriers, has considerable attraction. The protection enables the development of the "learning" (industrial) sector; the size provides scope for competition. (Our earlier remark explains why the degree of protection should be reduced over time.)

The structure of trade policy in the successful developing economies, like Japan, Europe after the Second World War, or other economies in Asia, has been very much of this sort. They have not focused on particular industries and protected them; they have tended to have broad protection across a range of industries, and they have actually encouraged competition behind those barriers.

The question of how this affects financial markets also arises—a question that Arrow's and Solow's work is particularly well positioned to help answer. When a country exports capital, the owners of that capital are, in effect, importing capital services from overseas. Just as imports of manufactured and industrial goods fail to carry with them the learning that is associated with those sectors, imports of financial services fail to carry with them the important learning that is associated with that sector. If there are powerful arguments for broad barriers to imported industrial goods, those apply equally to restrictions on capital exports overseas and the import of financial services. In short this theory provides a new rationale for why capital and financial market liberalization may lead to lower rates of growth. Similar arguments also apply, we show, to labor exports overseas.

From these analytics, the book moves to a broader policy discussion, beginning with trade and industrial policy, moving on to macro, financial, and investment policies, and to intellectual property. We explain why the political economy objection to specific infant-industry protection—that, for instance, the special interests that benefit from such protection work to keep it in place long after the economic justification for such protection has gone—have much less force in the context of the *infant-economy argument for protection*. We show that political economy concerns affect not whether there should be industrial and trade policies, but which policies and how they are best designed. We show too that intellectual property laws, if not well-designed, may actually impede learning and that "stronger" intellectual property regimes may be associated with a slower pace of innovation.

This part ends by moving beyond creating a learning *economy* to creating a learning *society*, and beyond the standard economic model, with its assumptions of rational individuals with predetermined preferences, incorporating insights from recent advances in behavioral economics, including the notion that preferences and beliefs are, at least in part, socially determined. We ask, for instance, whether there are policies that can help create a learning "mindset."

We hope this selection of insights has provided enough tempting morsels to persuade the reader to delve deeper into what follows. As we attempted to exposit our ideas, we faced a major dilemma: Mathematics is the language of modern economics. It can help ensure that putative conclusions follow from the assumptions. It can help test the robustness of the results: Do changes in assumptions lead to markedly different conclusions? But it can also obscure: the complexity of the analysis can also hide the role of particular assumptions. Arrow and Solow taught us the value of simple models—that we should strive to find the simplest and most general model to explore and explain the particular issue at hand. We hope the exposition here lives up to the high standards that they set.

But even the simplest analysis in this area can be relatively complex. And testing the robustness of the results requires exploring multiple variants of the basic model.

This book provides an opportunity to honor another of our teachers, Robert Solow, the father of modern growth theory. Solow and Arrow taught us how simple ideas can have profound effects. Bringing in insights from the economics of knowledge and learning fundamentally changes one's view about how to think about policies designed to promote growth. The infant-economy argument, inspired by Ken Arrow's paper on learning by doing, is, we believe, in the broad tradition of Ken Arrow and Bob Solow, in extending economic insights to new areas. We hope that the insights it provides will help poorer countries employ novel and effective policies to promote their economic growth and development.

CHAPTER ONE

The Learning Revolution

FROM ROMAN times, when the first data on per-capita output are available, until 1800, average human standards of living increased only imperceptibly if at all (see, e.g., Maddison 2001). Consumption for the great majority of human beings consisted predominantly of food, and food was largely limited to staples—rice, wheat, and other grains. Housing entailed barnlike living conditions with no privacy, and climate control consisted only of necessary heat in winter. Clothing was utilitarian and rarely involved more than single outfits with the seasonal addition of overclothes. Medical care was almost nonexistent. Travel was rare, largely local, difficult, and uncomfortable. Recreation was self-generated and primitive. Only a small aristocratic minority enjoyed what we would consider today an appropriate human standard of living—varieties of fresh food, including meat; private, well-warmed accommodations; multiple sets of clothing for varied occasions; rudimentary personal and medical care; and opportunities for travel and sophisticated entertainment.

Beginning in 1800 and accelerating markedly after the mid-to-late nineteenth century, that privileged standard of living began to diffuse throughout Europe, North America, and Australia. The impact of this change is apparent even in critical contemporary commentaries. The Communist Manifesto is in many ways a paean to the potential of the newly apparent economic progress—the benefits of which had not yet been widely shared.

In the twentieth century, elite standards of living became pervasive in Europe, North America, Australia, and many parts of Asia; a trend which continues in much of Asia today.

The significance of these transformations can be seen in another way: until the beginning of the nineteenth century, most individuals spent most of their time meeting the basic necessities of life—food, shelter, clothing. Today, for most of those in the advanced industrial countries—and for an increasing number in the emerging markets—satisfying these basic necessities of life takes but a few hours of work a week. Individuals can choose how to spend the "extra" time available: to work, to earn enough to consume more (whether higher quality "necessities" or luxuries) or to enjoy more leisure.[1]

What was the source of these societal transformations? Was it capital accumulation or technological progress? Although economists, such as Schumpeter (1943), had identified the major source of these transformative developments as technological progress, it was not until Robert Solow (1957) that there was a way of quantifying the relative importance of capital accumulation versus technical progress. Changes in capital intensity could account for at most a third of changes in output per worker. The remainder was attributable largely to various forms of technical progress.[2]

Subsequent literature suggested that the quantification was perhaps less robust than it seemed initially, partly because the measurement of key inputs (capital, human capital) was more difficult and problematic than had at first been realized, partly because the underlying model, entailing a constant returns to scale aggregate production function and full competition, seemed more questionable.[3] Some of the difficulties of parsing out the sources of growth was that they were intertwined—new machines (investment) were required to implement new technologies.[4] Still, there is no doubt that there have been enormous increases in productivity and that advances in technology as well as "learning to do things better" have played a critical role in these increases in productivity. For our purposes, that is all that matters.[5]

Not only is the pace of learning (innovation) the most important determinant of increases in standards of living, the pace itself is almost surely partially, if not largely, endogenous. The speed of progress has differed markedly both over time and across countries, and while we may not be able to explain all of this variation, it is clear that government policies have played a role. Learning is affected by the economic and social environment and the structure of the economy, as well as public and private investments in research and education. The fact that there are high correlations in productivity increases across industries, firms, and functions within firms suggests that there may be common factors (environmental factors, public investments) that have systemic effects or that there may be important spillovers from one learner/innovator to others. But the fact that there are large, persistent differences across countries and firms—at the microeconomic level, large discrepancies between best, average, and worst practices—implies that knowledge does not necessarily move smoothly either across borders or over firm boundaries.

All of this highlights that one of the objectives of economic policy should be to create economic policies and structures that enhance both learning and learning spillovers; creating a learning society is more likely to increase standards of living than is making small, one time improvements in economic efficiency or sacrificing consumption today to deepen capital.[6]

And this is even more so for developing countries. Much of the difference in per capita income between these countries and the more advanced is attributable to differences in knowledge. Policies that transformed their economies and societies into "learning societies" would enable them to close the gap in knowledge, with marked increases in incomes.[7] Development entails learning how to learn (Stiglitz 1987c).

Solow, in his seminal paper on the economics of growth (1956), had, for simplicity, modeled the rate of technological progress as fixed and exogenous, unaffected by the decisions of firms. This left unexplained the most important source of increases in living standards—and thus provided little guidance on how economic policy might increase that pace. Thus, Solow's 1957 paper showed that what his 1956 paper focused on, capital accumulation, was relatively unimportant; what was important was what his 1956 paper took as simply *given*. Not surprisingly, soon after Solow's pioneering work, there developed a large literature in growth theory attempting to "endogenize" technological change—starting at least as early as the 1960s,[8] with further progress being made during the 1980s.[9]

The best work tried, of course, to base the analysis of aggregate (macro) behavior on micro-foundations. There is, by now, a large literature on the microeconomics of technological progress,[10] but many of the insights of that literature have not been incorporated into the macroeconomic growth models, which often take a simplistic view, ignoring, for instance, sectoral differences in the pace of innovation, the multitude of ways in which progress occurs, and the interrelationships among them and alternative policies. To deal with the complexities posed by endogenous growth, and the challenge of deriving long-run steady-state growth, much of the literature has focused on parameterizations that turn out to be very, very special. While some of the literature has recognized that when innovation is endogenous, markets are not likely to be fully competitive, the interplay between market structure and innovation is typically not at the center of discussion. Is even the kind of competition that Schumpeter envisaged really viable? Some of the literature makes assumptions that virtually prejudge the conclusions: If trade is assumed to enhance learning (and more effectively than a corresponding amount of domestic production), then trade barriers have an adverse effect on economic growth. As we show, alternative (and we would argue more plausible) assumptions about the innovative process suggest that some trade restrictions may be desirable.

If our contention that the success of modern economies is due to innovation and learning is correct, then understanding the processes of learning and innovation, and how policy can affect its pace, should be at the center of economic analysis.[11] We can think of an economy's "innovation system" broadly as running from basic research—typically financed by government, occasionally by a government sanctioned monopoly (like Bell Labs), and typically produced by research universities and government research laboratories—to applied research, sometimes building on these basic ideas, at other times refining and developing "prior art." Ideas have to be disseminated and put into practice: much of the increase in productivity occurs as firms learn from each other or as technology improves through practice. More of our analysis ought to focus on how such learning occurs.

Kenneth Arrow was a pioneer in examining the economics of these "learning processes"—the factors which promote or retard them, their likely response to normal market incentives, and their relationship to the broader macro- and microeconomic environment—notably in his papers on the economics of R & D and learning-by-doing (1962a,1962b). He called attention to the fact that while some knowledge was produced as a result of the deliberate allocation of resources to research and development, much of technical progress was a by-product of production or investment.

One of the advances in modern economies has been improvements in the processes by which they learn—they have learned how to learn. There is not a single breakthrough that led to enhanced learning capacities, but rather a series of organizational innovations.

Consistent with this, subsequent work, including that of Nordhaus (1969a, 1969b), identified the greater part of such progress as arising from the continuous accumulation of small improvements in production processes rather than from dramatic technological breakthroughs, though some, perhaps many, of these small improvements may be based on or related to transformative changes. For example, computerization and electrification were big changes, but their full effects were brought about in small steps.[12] So too, the separation between capital accumulation and "learning" is not a clean one: It is often through new investments that new ideas are discovered and new research is "embodied."[13] If the pace of investment determines the pace of learning, then of course it is impossible to neatly separate out what part of the increase in productivity is a result of capital accumulation and what part is a result of improvements in technology, because the two are inextricably intertwined.

The highly aggregate models that have been at the center of modern growth and development theory miss another key point: In the standard paradigm, except for market distortions (and the elimination of such market distortions is the passion of most economists), firms are always assumed to be on the production possibilities curve (in the jargon of traditional economics). Productivity increases, in this standard model, result from moving the production possibilities curve out, as a result either of more accumulation of human or physical capital or of R & D. Indeed, much of the literature treated knowledge essentially as another form of capital—"knowledge capital"—ignoring its distinctive properties, which will be the focus of discussion in later chapters, especially chapter 6. But in reality, most firms operate well below their production possibilities curve. There are large gaps between "best practices" and "average practices." Countries differ in the size of these gaps. Closing these gaps can, at least for a while, provide an important impetus to societal increases in productivity. And for the typical firm, even as it closes yesterday's gap, new gaps open up. Most firms are forever "catching up."

The most successful economies are those that have succeeded in not only moving out their production possibilities curve more rapidly but also ensuring that the gap between "average" and "best" practices is small. There is more diffusion of knowledge, more learning; it is these achievements in learning that largely account for the ever-rising standards of living in these successful economies.

In short, *the transformation to "learning societies," which occurred around 1800 for Western economies, and more recently for those in Asia, appears to have had a greater impact on human well-being than improvements in allocative efficiency or resource accumulation*. If this is so, understanding how to create a learning society should be one of the central preoccupations of economists and other social scientists. Success in this endeavor can have a far greater impact on increasing long-term living standards than ascertaining how to increase resource accumulation or reduce short-term allocative inefficiencies.

This book seeks to present the simplest framework for understanding some of the critical determinants of the rate of progress—disaggregated enough for sectoral policies to make a difference, but aggregated enough to retain our focus on the determinants of the economy's overall rate of progress

Central to our enquiry are two basic questions: Do markets, by themselves, result in an efficient level and pattern of learning and innovation? And if not, what are desirable interventions by government?

1. Market Inefficiency

The answer to the first question is simple and straightforward: *There is no presumption that markets are efficient in the production and dissemination of knowledge and learning. Quite the contrary, there is a presumption that markets are not efficient*.

Modern notions of the efficiency of markets date back to the work of Adam Smith (1776) and his invisible hand: the notion that the pursuit of self-interest would lead, as if by an invisible hand, to the well-being of society. It would take 175 years before Arrow (1951b) and Debreu (1959) would establish the sense in which that was true (markets are "Pareto efficient"; that is, no one could be made better off without making anyone worse off) and the conditions under which it was true. Arrow provided *sufficient conditions* for the Pareto efficiency of markets (see Arrow 1951b; Debreu 1959). Subsequent work showed that those conditions were also essentially necessary. For instance, his proof of the efficiency of markets required that information was exogenous (that is, it needn't be perfect, but beliefs couldn't change as a result of what individuals observed or did); it was subsequently shown that whenever markets were incomplete or information was endogenous and asymmetric (that is, essentially *always*), markets were not (constrained) Pareto efficient.[14]

For the purposes of this book, though, the central assumptions in the proof of the efficiency of the market economy were that markets were perfectly competitive and that the state of technology was fixed, exogenous. Arrow and Debreu, in their proof of the efficiency of the market economy, assumed away innovation. In doing so, they left unanswered the question of whether a market economy was efficient in innovation. Given that many advocates of markets assumed that their innovativeness was their central virtue, this was obviously a central lacuna. Indeed, earlier Schumpeter (1943) had gone so far as to argue that one of the distortions on which many economists had focused attention—monopoly—could actually be a virtue in an innovation economy: it provided the rents which supported R & D, and so long as there was competition *for the market*, one should not worry about competition *in the market*. But neither Schumpeter nor others arguing for the virtues of markets on the basis of their innovativeness were able to show that markets were efficient in innovation.

The reason that they did not do so is that they could not do so: the discussion that follows, building on the work of Arrow and others, shows that there is a presumption that markets on their own are efficient neither in the level nor the pattern of innovation. Arrow recognized that market failures in the production and dissemination of knowledge (whether as a result of the allocation of resources to R & D or as a result of learning) were pervasive. Thus, following Arrow's lead in understanding the economics of learning processes—and the pervasive market failures in learning processes—is critical to formulating effective economic policies.

Analyzing the nature of these inefficiencies—and their implications for public policy—requires the construction of a general equilibrium model in which R & D or learning and market structures are both endogenous. The market inefficiencies are multiple and complex. We will explain, for instance, why some sectors may be more amenable to learning than others; why some sectors may generate more externalities (spillovers to other sectors) than others. We will see that sectors in which learning is important are often imperfectly competitive, so that not only may production—and learning—be constrained below its optimal level because firms fail to take into account the spillovers that their learning has for other sectors, but production—and learning—may be constrained as a result of the exercise of market power. We will explain why Schumpeter's view that such market power was of positive value (it helped finance research that otherwise would not have received funds) and that abuses would be limited because of the discipline of "Schumpeterian" competition (competition to be the dominant firm through innovation) needs to be qualified. His view on monopoly was too panglossian.

Further inefficiencies in the innovation process are introduced as a result of capital market imperfections and imperfections in risk markets. The Arrow-Debreu analysis establishing the efficiency of markets required not only unreasonable assumptions about the nature of competition and innovation but also that there be a complete set of risk markets and perfect capital markets. The imperfections in these markets, especially as they relate to innovation, are not just a happenstance but an inherent feature of innovation, as we explain in chapter 6.

2. The Role of Government in Promoting a Learning Society

If learning, and R & D more generally, is at the center of the success of an economy, and if there is no presumption that markets are efficient in making decisions which affect the pace of learning (or R & D), then long-standing presumptions against government intervention are simply wrong. The financial crisis has called attention to the role of government in crisis prevention. Widespread environmental problems have called attention to the role of government in preventing pollution and potentially catastrophic climate change. These are examples of government's role in preventing *negative* externalities. The production of knowledge entails *positive* externalities. The private sector produces too much of goods that give rise to negative externalities, which is why government must either impose charges when firms generate pollution or otherwise regulate pollution-generating activities. By contrast, the private sector typically produces too little of goods that give rise to positive externalities. Again, to correct this market distortion requires some form of government intervention.

These interventions, though, are more complex than those that are necessary to correct the negative externalities; there are limited and well-identified environmental externalities and a well-developed set of tools for addressing these market failures. So too there is widespread understanding of the externalities that can be generated by underregulated financial markets and, especially in the aftermath of the financial crisis, even an understanding of what good regulation entails. But learning touches every aspect of a modern dynamic economy, and even more so of an emerging market struggling to become an advanced industrial country. If there are market failures in learning, then the market failures are pervasive in the economy. They are diffuse. More pervasive government interventions are required to correct them.

Many of these advances on which our dynamic economy is based rest on government-funded research, and without that support, the pace of innovation—and the pace of increases in standards of living—would have been far lower. And many of the advances attributable to the private sector are shaped by our legal framework, including those governing intellectual property. Critics on both the left and right assert, though, that this legal framework may be far from ideal, some suggesting that innovation is impeded as a result of insufficient protection of property rights, others that progress is hampered because of a poorly designed intellectual property regime, more focused on increasing the rents of, say, the pharmaceutical industry than advancing standards of living. Whatever one's views on these questions, there is a consensus that public policy is central and unavoidable. Government has a responsibility in "creating a learning society." If we are to understand what that responsibility is—and how it can best be fulfilled—we have to understand why it is that markets on their own don't "work" and how innovation actually occurs in our society.

The analysis presented here thus changes the presumption about the desirability of government intervention: now there is a presumption of market failure and a presumption for government to take actions to correct these market failures.

This book, then, is an attempt to study the economics of "learning societies," focusing especially on the role of government in promoting growth through the creation or strengthening of a learning society. This book lays out simple models in which learning spillovers are well identified. The models generate policy prescriptions which differ markedly from standard policy recommendations focusing on enhancing allocative efficiency. It is not just a difference in emphasis between classical economic policy prescriptions—based on notions of static allocative efficiency and the idea that growth in productivity arises chiefly from resource accumulation (physical, human, and scientific capital)—and those that we stress for creating dynamic learning environments. Rather, our concern is that some of these classical policy prescriptions, though well-intentioned, may actually lead to a reduction in the rate of progress of societies and a deterioration in long-run societal well-being. In the attempt to improve the static efficiency of the economy, learning may be impeded. Our analysis supports numerous policies that have been proscribed by economists wedded to the neoclassical model and suggests new measures that will help create a more dynamic learning economy. In that sense, our work is similar to that of Schumpeter (1943), who criticized conventional economists for overemphasizing competition. But while Schumpeter was correct in his critique of neoclassical economics, he never formulated a coherent analytic normative or positive theory. The result is that some of his normative stances were misguided. For instance, he was overly optimistic (as we shall see) about the potential for what has come to be called "Schumpeterian competition" to ensure, by itself, a dynamic economy, and he was overly sanguine about the virtues of (temporary) monopolies.

This reassessment of policy is especially important for developing countries and emerging markets. As we noted earlier, what separates developed from less-developed countries is not just a gap in resources but a gap in knowledge. Thus, a central focus of development policy should be closing that gap—and that means enhancing learning. This is, for instance, one of the central objectives of modern industrial policies, which seek to promote particular industries and particular technologies with greater learning capabilities and greater spillovers to other sectors. (While industrial policies were originally targeted toward supporting the industrial sector, today the term is used much more broadly, to embrace any set of policies designed to encourage particular sectors or technologies. Policies promoting the agricultural sector, the research sector, or the service sector would thus be included under the term *industrial policy*.)[15] Policies that in any way impede learning—including those that seek to circumscribe the use of industrial policies—may, over the long run, lower well-being.[16] This is but one example of many places where we argue that traditional developmental policy stances, such as those associated with the Washington consensus, are misguided: Well-designed trade restrictions, subsidies, and exchange rate interventions can play an important role in promoting learning. We argue, furthermore, that removing domestic content restrictions on foreign direct investment, as called for by trade and investment agreements, may impede learning.

One of the clearest points of departure between our focus and that of more traditional development economics concerns the role of institutions. Much of the standard literature has emphasized the role of institutions that protect *property rights*. As knowledge has become more important, there is an increasing emphasis on intellectual property rights and the institutions that protect them. By contrast, we take a broader view: Intellectual property rights is one institution that incentivizes innovation. But there are others that are as or more important. We ask what the institutions are that promote a learning society. We also argue for an intellectual property rights regime that is markedly different from the regulations incorporated into the TRIPS agreement of the WTO. Indeed, we argue that poorly designed "strong" intellectual property regimes actually impede learning and innovation.

There are, in fact, many examples where the approach we take leads to policy recommendations that are contrary to those of the Washington consensus: We argue against measures on financial market liberalization that have typically been included in agreements signed under bilateral trade agreements and under the Financial Services Agreement of the WTO. We provide an explanation for why trade and capital market liberalization have often failed to promote growth in the way that was hoped, and we suggest how these measures should be modified, once one takes a learning perspective.

Much of this book centers around the question of how to best promote learning, including how to balance optimally the dynamic gains from faster learning with the short-run (static) costs associated with interventions, and how best to design interventions. But much of the debate around government intervention has centered on political economy concerns. These should not, and cannot, be ignored. We will argue that these have more to do with the form of government intervention than they do with *whether* there should be government intervention.

3. The Theory of Comparative Advantage Redefined

Perhaps the most important way in which our book differs from classical prescriptions is that we argue that there is an *infant economy argument for protection*. Growth and standards of living can be raised by *defying* a country's seeming comparative advantage and imposing trade restrictions that encourage industrialization. But our book also provides a different perspective on what is meant by comparative advantage. The traditional theory of comparative advantage (as developed by Eli Hecksher and Bertil Ohlin[17]), based on the notion that knowledge was fully available, focused on relative factor endowments. Portugal exported wine because it was endowed with weather more suitable for growing wine, England cloth. Countries that had an abundance of unskilled labor exported unskilled labor intensive goods.

Krugman's (1979) research building on the Dixit-Stiglitz model of product differentiation made it clear that something besides factor endowments mattered. He observed that most trade today is between countries that have similar factor endowments. And he observed that they often traded similar products. Germany exports cars to the United States, and the United States exports cars to Germany and other countries. But in the Krugman-Dixit-Stiglitz model, there is no explanation of why Germany is exporting the kinds of cars that it does. There are multiple equilibriums: the United States could have wound up exporting the cars that Germany did and vice versa. Our analysis suggests that, to a large extent, these patterns are not just the outcome of fortune, the toss of a coin, but are related to the more fundamental endowments—the state of knowledge and learning capabilities.

Justin Lin (2012) has distinguished between industrial policies that defy comparative advantage, which he argues are likely to be unsuccessful, and those that are consistent with comparative advantage, which can be an important component of successful development. While there is considerable insight in this distinction, the key question is, what are a country's *endowments*, which determine its comparative advantage? This is equivalent to asking, what are the relevant *state* variables, those that describe the state of the economy today? And what is the "ecology" against which the country's endowments are to be compared, i.e., what are the *relevant* endowments of other countries?

It has become conventional wisdom to emphasize that what matters is not static comparative advantage but dynamic comparative advantage. Korea did not have a comparative advantage in producing semiconductors when it embarked on its transition. Its static comparative advantage was in the production of rice. Had it followed its static comparative advantage (as many neoclassical economists had recommended), then that might still be its comparative advantage; it might be the best rice grower in the world, but it would still be poor. But a country's dynamic comparative advantage is endogenous, a result of what it does. There seems to be a circularity here. The central question is, what should a country do today to create its dynamic comparative advantage?

Ascertaining a country's static comparative advantage is difficult; ascertaining its dynamic comparative advantage is even harder. As we noted, standard comparative advantage focused on *factor* endowments (capital-labor ratios). But with capital highly mobile, capital endowments should matter little for determining even static comparative advantage. Still, capital, or more accurately, the knowledge of the various factors that affect returns and that is required to use capital efficiently, doesn't move perfectly across borders; neither does knowledge about how effective a particular enterprise is in using various inputs to produce and market outputs. That means that the resident of country *j* may demand a higher return for investing in country *i* than they would demand for investing in their own country. There is, in practice, far less than perfect capital mobility.

The "state" variables that determine comparative advantage relate to those "factors" that are not mobile, which, in varying degrees, include knowledge, labor, and institutions.

Multinationals can, however, convey knowledge across borders. Highly skilled people move too. Migration has resulted in large movements in unskilled labor but, in most cases, not enough to change endowments of the home or host country significantly. Even institutions can sometimes effectively move across borders, as when parties to a contract may agree that disputes will be adjudicated in London and under British law. Still, there are numerous aspects of tacit knowledge, about how individuals and organizations interact with each other, and norms of behavior that affect economic performance and, most particularly from our perspective, how (and whether) they learn and adapt. Such tacit knowledge does not typically move easily across borders. (Indeed, as we argue, it does not even move easily among or within firms. There are natural barriers to the flow of knowledge, including incentives on the part of market participants to take actions that would impede the flow of knowledge.)

The most important "endowment," from our perspective, is a society's learning capacities (which in turn is affected by the knowledge that it has; its knowledge about learning itself; and its knowledge about its own learning capacities), which may be specific to learning about some things rather than others. The spirit of this book is that a country's policies have to be shaped to take advantage of its comparative advantage in knowledge and learning abilities, including its ability to learn and to learn to learn, in relation to its competitors, and to help develop those capacities and capabilities further. Even if it has the capacity to learn how to make computer chips, if a country's learning capacity is less than its competitors, it will fall behind in the race. But each country makes, effectively, decisions regarding what it will learn about. There are natural nonconvexities in learning, benefits to specialization. If a country decides to learn about producing chips, it is less likely that it will learn about some other things. There will be some spillovers to closely related technologies—perhaps to, say, nanotechnology. The areas to which there are spillovers may not lie nearby in conventional product space. There may, for instance, be similarities in production technologies (as in the case of just-in-time production or the assembly line). That is why the evolution of comparative advantage may be so hard to predict.

But while standard economic analysis may provide guidance to a country about its current (static) comparative advantage (e.g., given current technology, for a country with an abundance of unskilled labor, what are the unskilled-labor intensive goods), guidance about its comparative advantage defined in this way (dynamic learning capacities) is much more difficult. In part, this additional difficulty is because this advantage depends on judgments by other countries about their dynamic comparative advantage and their willingness to invest resources to enhance those advantages. Whether ex ante the United States, Japan, or Korea initially had a dynamic comparative advantage in producing chips, once Korea had invested enough in learning about certain kinds of chip production, it would be difficult for another country to displace it. Another country would have to leapfrog—and whether it could do so depends not only on the other country's capabilities and its willingness to invest to enhance those capabilities but also on Korea's responses to these competitive threats.[18]

Looking at what other countries at similar levels of per capita income did in the past or what countries with slightly higher levels of per capita income are doing today (as Lin suggests) may be helpful, but only to a limited extent. The world today (both in terms of global geo-economics and geo-politics, and technology) is different than it was in the past. Competing in textiles today requires different skills and knowledge than in even the recent past; a lagging country wanting to enter a market may (or may not) be able to displace a country that currently has a comparative advantage in some product; that country may (or may not) be in the process of attempting to establish a comparative advantage in some other area.

In short, the learning perspective redefines the theory of dynamic comparative advantage and does so in ways which make formulating development strategies more complicated—but more interesting. Less-developed countries today cannot simply imitate patterns of development that were pursued by earlier developers. That this is so should be evident by now. Those countries in the early to mid-twentieth century that followed the heavy industrialization strategy that was the basis of the success of the United States and Germany in the nineteenth century failed. African countries that try to follow blindly the export-led strategies of East Asia may find them far less successful than they were when they were employed in East Asia in the latter third of the twentieth century. While development economists are likely to praise East Asia's export-led growth strategy, it was not growth in exports per se that led to their success; it was growth in particular kinds of exports that were associated with high levels of learning. Other countries pursuing export-led growth strategies but exporting goods for which there are not such learning benefits may find themselves sorely disappointed.

This discussion highlights the important ways in which the learning perspective redefines basic concepts, like comparative advantage, policies, and economic strategies. The learning perspective also leads to rethinking other long-standing notions. We have already noted that our theory calls into question the usefulness of the concept of an aggregate production function, especially one predicated on the assumption that all firms (say, within the country) have the same knowledge and are equally capable of converting inputs into outputs. In chapter 2 we reconsider the concept of the production possibilities schedule, and in chapter 4 we consider the question, posed some 75 years ago by Ronald Coase, of the boundary of the firm—what activities go inside the firm and what goods and services are purchased in the market.

CHAPTER TWO

On the Importance of Learning

THE CENTRAL thesis advanced in chapter 1 was that what distinguishes the modern era of the last two hundred years from the millennia that preceded is *learning*—we have learned how to increase productivity, the outputs that can be produced with any inputs. There are two aspects of learning that we can distinguish: an improvement in best practices, reflected in increases in productivity of firms that marshal all available knowledge and technology, and improvements in the productivity of firms as they catch up to best practices. In fact, the distinction may be somewhat artificial; there may be no firm that has employed best practices in every aspect of its activities. One firm may be catching up with another in some dimension, but the second firm may be catching up with the first in others. In developing countries, almost all firms may be catching up with global best practices; but the real difference between developing and developed countries is the larger fraction of firms that are significantly below global best practices and the larger gap between their productivity and that of the best-performing firms.

While we are concerned in this book with both aspects of learning, it is especially the learning associated with catching up that we believe has been given short shrift in the economics literature, and which is central to improvements in standards of living, especially in developing countries. But as we noted in chapter 1, the two are closely related; because of the improvements in best practices by the most innovative firms, most other firms are always engaged in a process of catching up.

While the evidence of Solow and the work that followed demonstrated (what to many seems obvious) the importance of learning for increases in standards of living, to further explicate the role of learning, the first three sections of this chapter marshal other macro- and microeconomic evidence. In particular, we stress the pervasive gap between best practices and the productivity of most firms. We argue that this gap is far more important than the traditional allocative inefficiencies upon which most of economics has focused and is related to learning—or more accurately, the lack of learning.

The final section provides a theoretical context within which to think about the sources of sustained increases in standards of living, employing the familiar distinction of movements of the production possibilities curve and movements toward the production possibilities curve. Using this framework, we explain why it is that we ascribe such importance to learning.

1. Macroeconomic Perspectives

There are several empirical arguments that can be brought to bear to support our conclusion concerning the importance of learning. The first is a simple argument: In theory, leading-edge technology is globally available. Thus, with sufficient capital and trained labor (or sufficient mobility for capital and trained labor), all countries should enjoy comparable standards of living. The only difference would be the rents associated with *ownership* of intellectual property rights and factor supplies. Yet there is an enormous divergence in economic performance and standards of living across national economies, far greater than can be explained by differences in factor supplies.[1] And this includes many low-performing economies with high levels of capital intensity (especially among formerly socialist economies) and highly trained labor forces. Table 2.1 presents a comparison of formerly socialist countries with similar nonsocialist economies in the immediate aftermath of the collapse of the state-controlled model of economic activity.

TABLE 2.1
Quality of Life Comparisons, 1992–1994 (U.S. $)

	Formerly Centralized			Noncentralized		
Region	Country	Per Capita GDP	Life Expectancy	Country	Per Capita GDP	Life Expectancy
Baltic	Lithuania, Latvia, Estonia	7,800	70.4	Finland	16,150	76.2
Middle Europe	Czech Republic	7,350	73.5	Austria	17,500	76.9
Agricultural Europe	Poland	4,920	73.1	Spain	13,125	77.9
China	China	2,500	68.1	Taiwan	12,070	75.5
Southeast Asia		870	55.3	Thailand	5,970	68.4
Korea	North Korea	920	70.1	South Korea	11,270	70.9
Unweighted Average		4,060	68.4		12,681	74.3

Source: Greenwald and Khan (2009), p. 30.

In most of these cases, at the time communism was imposed after World War II, the subsequently socialist economies enjoyed higher levels of economic development than those of the "comparator." Czechoslovakia was more highly industrialized than Austria. Finland was perhaps the poorest of the Baltic countries. Spain, a large Catholic agricultural country, was poorer than Poland. Taiwan, occupied by the Japanese for many years, was a relatively backward part of China. Vietnam and Cambodia were at least as well-off as Thailand. And North Korea was more heavily industrialized than South Korea. Over the intervening 40 years from the late 1940s to the late 1980s, the socialist economies focused strongly on the traditional growth prescriptions of capital accumulation and education. They had high savings and investment rates—far higher in many cases than those in the West—and even invested heavily in education, especially the kinds of technical education that might seem most directly relevant to production (and even some forms of innovation). Yet by the end of that period, they had levels of economic output less than one-half of that of the comparison economies (and often far less than one-half).

On one side of the line, economies developed based on steady improvements in economic performance over time. On the other side, economies largely failed to "learn," even if they did better in accumulating factors of production and even, in some cases, if they did better at developing advanced products, like the Sputnik. These countries (and the firms within them) not only failed to make productivity-enhancing innovations, they failed to learn from the innovations and best practices that were going on in the other parts of the world.

It should be clear that the differences that emerged were beyond those that could be explained simply by static inefficiencies (e.g., those associated with distorted incentive systems and the misallocation of resources). If that had been the key problem, then the move from communism to a market economy would have quickly closed the gap; moving to market prices and incentive structures should have eliminated these static inefficiencies. In fact, in most countries of the former Soviet Union, output actually fell (see Stiglitz 2000c). This is not to say that, for instance, distorted incentive structures played no role. When China moved from collective farming to the individual responsibility system, there was a large increase in productivity; but even then, productivity remained markedly below that of other countries. The magnitude of the gaps and their evolution over time (both before and after the end of communism) suggest that they cannot simply be attributed to static inefficiencies.

The same argument is, of course, true for changes over time. The same changes are available globally—especially for the many aspects of technology not protected by intellectual property—yet there are large differences in changes in productivity, differences not accounted for by differences in changes in other factor inputs.[2] This is again illustrated by the economies in transition. Indeed, nothing could illustrate more the significance of learning and learning capacities than the experience of those (and other) economies since the transition. Movements to higher standards of economic performance after the transition have been far from uniform. Some countries adapted quickly and well. From 1975 to 1980, reported annual per capita income growth in China was 4.1 percent. From 1980 to 1985, after altering the conditions under which businesses could operate and learn, growth accelerated to 8.4 percent, and since 1985 has been about 10 percent per year.

This rapid turnaround could not be attributable either to education or to capital accumulation.[3] A reformed education system would take at least 8 years before it could produce more highly trained graduates (since older classes would be inadequately prepared by their prereform training), and these graduates would transform the total labor force only slowly over time.[4] As for capital accumulation, even if the fraction of GDP devoted to investment were to have increased by 25 percent, at a real return of 5 percent, the acceleration in growth would have been just 1.25 percent; if the real return were even 10 percent, the acceleration of growth would have been just 2.5 percent. Clearly what changed was the effectiveness with which capital and labor were being employed using technologies that were preexistent and widely available globally.

Of course, improved incentives[5] in, say, agriculture and the reduction of sectoral misallocation of resources played some role in China's growth. Improvements in productivity resulting from the removal of a static inefficiency result, as we discuss more thoroughly later, in a one-time (or short-lived) increase in productivity were important. But while that was true (to a large extent) in agriculture,[6] it is striking that the most important sources of growth were in manufacturing, where one could visibly see improvements in productivity, quality, and practices, and these improvements were persistent.

Other formerly socialist countries experienced much slower convergence to high performance levels of per capita income.[7] In general the

Baltic countries and many of those in Eastern Europe "learned" far more slowly than their Asian counterparts (at least as reflected in rates of increase in aggregate productivity[8]), China and Vietnam. When they did eventually start to grow more rapidly, it was often through a real estate bubble: global financial practices—good and bad—seem among the easiest to move across borders.

India had a similar experience of accelerating growth after commercial reforms in the 1980s. It was these reforms, not the later trade liberalizations, from which one dates the rapid changes in India's growth (see Rodrik and Subramanian 2005).

Other countries, both formerly socialist and in Latin America, Africa, and parts of Asia, have yet to experience such high rates of growth. This has been true despite their frequent embrace of accepted market principles, their access to global technology, high rates of savings, sound macroeconomic policies, and well-developed educational systems. What has failed in these countries is, to a large extent, their ability to adapt existing global technology and deploy resources effectively *within each sector*. They remain mired well inside their *theoretical* production possibility frontiers.

2. Microeconomic Perspectives

Even in highly developed economies, like the United States or Japan, there is substantial evidence (see, e.g., Baily and Solow 2001) that most firms operate well below their theoretical capabilities (the "best practices" within the industry), implying a large scope in productivity increases from movements to the production possibility frontier by each firm.[9] Although, clearly, unrealized potential gains would eventually be exhausted without leading-edge research and development, for practical purposes, "learning" to exploit existing opportunities and the diffusion of existing technology contribute more to rates of productivity growth *at any particular moment* than leading-edge technological improvements.

One of the most striking aspects of firm-level studies of productivity is the existence of large and persistent productivity differences across firms, both at the level of overall output and at the level of the individual processes that generate overall output. At the firm level, differences in productivity of two-to-one or more between leading firms in an industry

TABLE 2.2
Data—Life Insurance Companies (General Expense as Percentage of Premiums)

Year	Connecticut Mutual	Phoenix Mutual	Northwestern Mutual
1988	20.9	16.7 (17.6)	6.8
1989	19.8	15.7	6.9
1990	20.2	14.9	7.4
1991	20.9	15.6 (18.2)	6.3

and the industry average were first documented systematically by Baily et al. (1992) and have been confirmed by most subsequent studies.[10]

An example for mutual life insurance underwriters is presented in table 2.2. Northwestern Mutual, the acknowledged industry leader, was able to process life insurance premiums at a cost of about 40 percent of that of an average performer, like Phoenix Mutual, and less than a third of that of a relative laggard like Connecticut Mutual. Correcting for differences in operations, like product mix and organizational form (e.g., using proprietary sales agents versus independent agents), actually increased Northwestern's measured performance advantage. For example, Northwestern Mutual sold a higher proportion of term-life policies (as opposed to whole-life policies) than its competitors, and term-life policies with lower premiums per policy typically required a higher administrative effort per premium dollar than whole-life policies.

Just as strikingly, another feature of table 2.2 appears to be broadly representative of within-industry performance differences: There is only limited convergence in productivity levels between leading companies and their less efficient competitors. (An equilibrium model explaining why that might be so was presented in the appendix to this chapter that was included in the unabridged version of this book.)

Baily et al. (1992) and others (see, e.g., Dwyer 1998) typically find that rates of convergence for productivity levels across firms within an industry are very slow indeed. Leading firms with successful learning environments appear to increase productivity at rates which are comparable to those of their less efficient competitors despite being necessarily nearer the industry production possibility frontier. These firms seem more capable of learning.[11]

Of course, to remove redundancy, leading-edge firms are also likely to be operating well below what might be feasible, even with reasonable levels of investment, say, in engineering or new technology. This highlights a point of general relevance: the distinction between learning involved in moving toward the leading-edge technology and learning by those at the frontier may be less than is commonly thought. Moreover, even advances in leading-edge technology are typically the result of small improvements—not big innovations of the kind covered by the patent system. They are the result of learning—learning from doing and learning from others, figuring out what ideas and practices in other industries, for instance, are relevant to, or can be adapted to, the industry or enterprise in question.

The inescapable conclusion from this firm-level data is that most firms operate well within their industry production possibility frontier. But if firms operate inside their production possibility frontiers, then it follows that economies as a whole operate below their levels of optimal output. The potential for learning-driven output growth is clearly apparent in the microeconomic data.

3. Evidence from Episodes of Rapid Productivity Increases

The existence of this unexploited potential productivity is confirmed under some special historical circumstances where there was a sudden necessity to increase output. For example, a labor agreement in the U.K. engineering industry in the 1980s provided for a workweek reduction from five to four days at proportionately reduced wages. The idea was that employment would increase so that the available work would be spread among more union members. In response, process changes at the firms, forced to accommodate new working schedules, led to further significant reductions in employment despite increasing industry output.[12]

Another firm-level example of unexploited productive capacity can be seen on the occasion of a strike at the New York Telephone and New England Telephone companies in 1989. The firms had 80,000 workers prestrike; of these, 57,000 went on strike. During the first week of the strike, 22,000 of the 23,000 managers were assigned to cover for the missing union members. Their learning curve was so sharp that during the second week of the strike, half of these workers (11,000) were able to be reassigned to their original jobs and all the prior management work continued to be performed. The only normal work not being done during this second week was residential phone installations that involved rewiring the network and some new plant construction. Both functions could have been completely covered by hiring an additional 3,000 workers (using prestrike industry productivity norms). Under the special pressures of the strike, 26,000 workers fulfilled the role of 80,000 prestrike workers, a threefold increase in productivity. The evidence again argues for a substantial gap between where economies typically operate and the true frontier of potential production.

The fact that it is possible to increase productivity quickly—without dramatic changes in technology or inputs—provides further evidence for the potential role of learning.[13] If firms were indeed achieving their full productive potential, then further improvements in performance should be relatively slow, steady, and positive. Changes in the quality of a firm's labor force take place only slowly; most employees do not turn over during the course of a year, and new employees tend to have qualifications similar to existing ones. Capital additions during any particular year also tend to occur at a relatively stable rate and change the existing stock of capital only marginally. Finally, dramatic technological breakthroughs are rare, and firms most often adopt proven technologies rather than transformative leading-edge new ones.

In practice, however, productivity growth at the firm level tends to be highly episodic. The question is, why are such opportunities recurrently available? A firm on or near its production possibility frontier should not be able to achieve such sharp short-term improvements in operations (usually without significant investment or employee turnover), and it shouldn't be able to achieve such cost reductions repeatedly. All of this strongly suggests that productivity shifts at the firm level often, or even typically, consist far more of movements *toward* the production possibility frontier than of movement in the frontier itself.[14]

Macroeconomic Episodes

At the macroeconomic or sector level, there is also strong evidence of the disproportionate importance of "learning" environments. The most compelling example of this was the performance of the United States economy during World War II. Notwithstanding the massive shifts in production to war material and of manpower to the armed forces, output of consumption goods actually increased between 1941, by which time mass unemployment had largely disappeared, and 1945.

Another major example is suggested by the performance of the U.S. manufacturing sector between the 1970s and early 1980s on the one hand and the late 1980s and 1990s on the other. Between these two periods, the annual rate of growth of U.S. manufacturing productivity rose from 0.9 percent to 2.9 percent. The improvement coincided with a marked rise in U.S. real interest rates (normally associated with *less* investment in technology) and government deficits, a decline in U.S. research and development spending, and no detectable improvement in the performance of U.S. education (as measured by standardized tests). At the same time, it cannot be attributed to the availability of new technology. Such technology would have been equally available to other G7 economies. Over the period in question, the U.S. improvement in annual manufacturing productivity growth was 1.9 percent higher than that of the other G7 countries. The improvement was thus a U.S., not a global, phenomenon. What seems to have changed in U.S. manufacturing was an intensified focus on improved operations management through the rigorous implementation of procedures like benchmarking, total quality management, and reengineering—in our language, an intensified focus on learning. America seemed to have learned how to learn.[15]

Table 2.3 illustrates that something similar took place in the U.S. economy as a whole during subsequent years. After decades of productivity growth at rates well below those in Europe and Japan, U.S. productivity growth performance outstripped that of all these rivals in the years between 1995 and 2001. And again the relative changes involved were not related to changes in either capital accumulation (U.S. investment rates were little changed[16]), educational improvements, or formal R & D spending. They appear to have been rooted in improved learning in the United States.[17]

4. Alternative Theories of Growth

So far, we have presented convincing evidence that economies and firms operate well below the production possibilities frontier—what they could have produced, given the current state of knowledge, the best practices that are available within the economy—and that much of

TABLE 2.3
Productivity Growth Total, 1996–2001

Country	Change in Output per Worker (%)	Change in Hours per Worker (%)	Change in Output per Hour (%)
United States	11.4	−2.2	13.6
Canada	9.6	2.2*	7.4
Japan	6.4	−2.1	8.5
Germany	1.0	−8.5**	9.5
United Kingdom	7.2	−1.0	8.2
Italy	6.3	−0.3	6.6
France	5.2	−4.0	9.2

Sources: European Community Statistical Annual; U.S. Department of Commerce; U.S. Department of Labor; Canadian Government Statistics.

*Hours paid.

**Hours paid, major statistical revision in 2000.

the growth in productivity can be related to moving toward the frontier. While this is especially true in developing countries, it is even true in advanced industrial countries, highlighted by the marked differences among firms within the same industry in the same country.

We have suggested, moreover, that much of what occurs in this process of moving to the frontier can be described as "learning," catching up to best practices.

A standard analysis breaks down increases in productivity (output per worker, say) into two parts: How do we move economies to the frontier, and how do we move the frontier out? By the same token, policy analysis focuses on why the economy might be below the production possibilities curve (looking for reasons other than learning, with a focus on static inefficiencies), and why it might not be moving the frontier out in the optimal way. Such an analysis begins with the presumption that well-functioning economies operate at or near this frontier—as opposed to the presumption suggested in the first part of the chapter that even in a well-functioning economy, most or, indeed, even all firms operate well within the frontier of what is possible. In traditional approaches, learning plays little or no role, and, if it occurs, it is simply *exogenous*, that is, goes on independent of anything we do, how we structure the economy, how we restructure firms, and so forth. All firms have access to and make full use of all relevant knowledge.

In the traditional perspective, moreover, the only reason that firms would not operate at the frontier is if the government imposed distortionary taxes or regulations or did not prevent monopolies.[18] (There are a few other instances in which markets might "fail" to produce efficient outcomes, such as pollution, where one firm's pollution damaged another firm, but these too typically were given short shrift, particularly because they are easy to resolve, at least in principle, simply by imposing optimal corrective taxation [abstracting from the difficulties of politics]. Externalities mattered not so much for producers as for consumers, who might live a shorter life as a result.) A first task of policymakers, then, is to eliminate these sources of *allocative efficiency*.

However, dating back to the work of Harberger (1954), there has been a strong sense that the loss in welfare arising from these distortions is small. Hence these interventions, while beneficial, have impacts which are an order of magnitude smaller than the effects of movements in the frontier as well as those arising from macroeconomic disturbances, the periodic recessions and depressions that have plagued capitalism since its beginning and which leave large amounts of resources idle (so the economy is operating well below the production possibilities curve).[19]

Moreover, a movement toward the frontier results in a *one-time* increase in GDP, not a persistently higher level of growth. Even small increases in such growth rates—in the pace at which the frontier moves out—can, in the long run, lead to far larger increases in long-term GDP than the elimination of allocative inefficiencies.

Moving the Frontier Out by Investments in Capital and People

In the standard theory, then, sustained increases in standards of living, at least within developed countries, are largely associated with investments in capital and people that move the frontier outward. Within the United States, for instance, public discourse about the challenges of slow growth and the loss of international competitiveness have focused on increasing the amount and quality of physical and human capital. Concretely, discussions have centered on the quality of education (viewed as the major impediment to increases in human capital) and low levels of savings (the household savings rate reached near zero in the years before the 2008 crisis).

But Solow's brilliant 1956 paper laid to rest the view that higher savings and investment rates (including investments in human capital) would lead to sustained higher growth rates. He showed that a higher savings rate would lead to higher levels of per capita income but not permanently higher growth rates. The growth benefits might last longer, but the costs were more apparent than those associated with improving allocative efficiency. While the latter were typically described as (potential) Pareto improvements—in which *everyone*, now and in the future, could be made better off—the former entailed sacrificing current consumption for higher future consumption.

Hence, in the absence of a market failure, there is no presumption that individuals will save too little—that growth will be too low *even in the short run*. Whether increasing growth was socially desirable depended on intertemporal judgments—weighing the higher standards of living of future generations against the lower standard of living of the current generation. When, because of technological progress, future generations are likely to be much better off than the current generation in any case, it is not always convincing to ask further sacrifices of today's workers.

Conceptually, even with a low savings rate, there were two other ways of increasing GDP, the production that occurred within a country, within the standard paradigm. One was to import capital, though the benefits of the increased domestic output might accrue largely to the suppliers of capital. Within developing countries, there was much discussion of what they could do to attract more investment, from taking actions which ensured that the citizens of the country would receive ever less of the benefits of growth (e.g., as a result of tax holidays and land and other grants), to eliminating artificial impediments to the movement of foreign firms into the country. In some cases, so enthusiastic did governments become in recruiting investment that they "gave the company store away"; that is, while GDP went up—the value of what was produced inside the country increased—GNP, the income of the citizens, decreased. It is the latter (correctly measured) that matters, of course, and in many cases once the environmental and health effects were taken into account, the benefits to those living in the country were even more negative.

In these circumstances, it is not surprising that foreign direct investment in some areas can generate enormous opposition. Later in this book, we will explain why this neoclassical analysis leaves out one of the most important potential benefits—those associated with learning—but that these learning benefits won't necessarily come on their own: policies have to be designed to maximize them. And some international agreements are designed to limit the benefits that can accrue to developing countries. In some arenas, the *net* learning benefits may even be negative; that is, *from the perspective of learning*, given the constraints imposed by these international agreements, the country might be better off without (at least some of its) foreign direct investment.

There is a second way to improve the outward movement of the production possibilities curve with a given rate of savings, and that is to improve the allocation of capital—how it is deployed. In the standard neoclassical economy (which is serving as a foil to the learning economy that we analyze here), capital goods are allocated efficiently, or at least would be in the absence of government distortions. One could not obtain more growth from the given level of savings.

In short, we have argued that prospects for sustained increases in growth rates under the traditional neoclassical perspective are limited. There is a once-and-for-all gain as a result of eliminating static inefficiencies. If these are eliminated quickly, there is a rapid increase in growth while they are being eliminated, but then growth slows down to the rate at which the production possibilities schedule moves outward. Even the "growth" perspective focusing on increased savings has limited growth benefits, and these have to be offset by the significant costs associated with sacrifices in current consumption. Traditional analyses provide an uncompelling case for government intervention to increase the savings rate, and therefore even growth in the short run.

The one argument for government intervention that we have raised (beyond the direct role of government in enhancing societal learning, which is the focus of this book) is itself beyond the neoclassical model; it is that markets by themselves may be associated with inefficiencies in the allocation of capital and excessive instability. But even here, we argue that the standard analysis does not fully capture what is going on. Chapter 4 explains why instability is adverse to learning. There is thus a further long-run benefit to government policies directed at stabilizing the economy—not only do such policies result in fuller and more efficient utilization of resources, they may lead to systematically higher rates of productivity increase.

Reexamining Conceptual Foundations

Our analysis calls into question not only the traditional frame for parsing the sources of economic growth but even the underlying notion of a production possibilities schedule. If we assume that it represents the maximum level of output of, say, one product, given the output of others, *given the state of knowledge of each firm*, then it subsumes the knowledge of each market participant. It begs the key questions of why differences in knowledge persist, what can be done to reduce the gaps, or what limits the scale of production of, say, the more efficient firms.

Many years ago, Nicholas Kaldor similarly argued against the notion of a production possibilities schedule for a firm—and, by implication, for an economy, on a somewhat similar basis. He suggested that a firm typically had knowledge of its own production processes and knowledge of some nearby deviations. To be sure, there might be markedly different capital-labor ratios for which technologies *could* be developed, but such technologies did not exist and could only be brought into existence by investments in engineering (see Atkinson and Stiglitz 1969).

By the same token, the traditional approach encourages tautologies, such as treating differences in the knowledge of workers which might affect their productivity as differences in human capital. By definition, then, outward movements of the production possibilities curve as a result of learning are transformed into outward movements of the curve as a result of the accumulation of human capital. One might argue that this is a distinction without a difference; but there is a difference: Here (as we shall emphasize shortly), we want to understand the learning process. Are there, for instance, ways of organizing the work experience that accelerate the learning process? Is there something that can be done to enhance the learning capacities of individuals? What are the trade-offs? What does a firm, or a society, have to give up if it is to become more dynamic?

5. Concluding Comments

It follows from the analysis of this chapter that it is possible for countries to increase their rate of growth, if not permanently, at least for an extended period of time, well beyond that associated with a one-time improvement in allocative efficiency, or even beyond that associated with an increase in the savings rate. (In some of the models we construct later, we are able to show that countries can permanently increase their rate of growth.) The way they can do this is to create a learning society. For developed countries, this means ensuring that all firms learn quickly to improve their productivity, as best practices themselves improve, so that the gap between average and best practices is reduced; distorting resource allocations toward sectors with more learning and more learning spillovers; and investing more in R & D and in learning to learn. For developing countries, it means doing all of this, but with an eye mostly on closing the gap between their firms and best practices in the advanced industrial countries. Some resource allocations have more potential for doing so than others; the learning generated in some sectors and with some technologies has a greater potential of generating spillovers to others and enhancing societal learning capacities.

As we have emphasized, the fact that as firms move toward the frontier, the frontier itself has moved out means, of course, that the gap between best practices and average practices is never eliminated: learning is a perpetual process. And as we have suggested at a number of points, there may be less to the distinction between the two types of learning—moving the frontier out and moving closer to the frontier—for there is typically room for improvement in even the industry leader, and it may have something to learn from other firms both in its industry and in others.

This book focuses on learning as the basis of sustained growth and development, either to catch up with the best practices or to improve upon them. This chapter has presented micro- and macro-evidence supporting in particular the proposition that the productivity of firms and economies is well below what it might be, given the extant state of knowledge, and has highlighted the importance of learning in closing that gap. The appendix to this chapter (not shown here but included in the unabridged version of this book) constructs an equilibrium model in which it is optimal for some firms always to remain at some distance from best practices. There is always a knowledge gap. The next chapter takes a closer look at how learning takes place and the determinants of the pace of learning, the critical ingredients in creating a learning society.

CHAPTER THREE

A Learning Economy

A "LEARNING society" perspective takes a very different view of growth and development strategies from the standard neoclassical approach in several respects. It begins, as the last chapter explained, by focusing on the knowledge embedded in individuals, firms, and in society more generally—and how that knowledge changes, is transmitted, and is put to use. It recognizes that the state of knowledge of each individual in the economy can be (and typically is) markedly different. Knowledge, like information, is *asymmetric*. Each individual knows things that others don't. From the perspective of the individual, an advance in his or her knowledge is simply knowing something that he or she had not previously known. Indeed, in perhaps most cases, the knowledge that an individual obtains is already known (in some sense) by someone else; in a few cases, the knowledge that an individual obtains may not be known by anyone else, at least in a form which is easily recognizable. While the consequences of information asymmetry have been extensively explored over the past forty years,[1] the consequences of differential knowledge have not.[2]

Creating a dynamic learning society has many dimensions: individuals have to have a mindset and skills to learn. There has to be some motivation for learning. Knowledge is created by individuals, typically working within organizations, and transmitted to others within the organization. It is then transmitted from one organization and individual to another. But the extent, ease, and rapidity of transmission of knowledge is itself one of the central features of a learning society: for the new knowledge spurs new thinking; it is the catalyst as well as the input out of which new ideas and creativity emerge.

Some societies are better at learning than others—both in ensuring that the gap between best and average practices is smaller and in the pace with which the knowledge frontier moves out.

This chapter explores some of the elements that make for a learning society, asking how we can create an economic architecture that facilitates learning. Conventional discussions, especially among certain policy circles, focus too narrowly, for instance, in providing better incentives through greater appropriation of returns through stronger intellectual property rights (IPR). Not only is this view too narrow—as we explain later in this chapter, nonpecuniary intrinsic incentives may play a more important role in motivating learning than IPR—but it may be misguided, with stronger IPR encouraging secrecy and impeding the transmission of knowledge. As we showed in appendix A to chapter 5 of the unabridged version of this book, the result may be that innovators take more out of the pool of opportunities than they contribute, resulting in a smaller pool of opportunities and a slower pace of innovation. This book will explain how many other aspects of the legal, institutional, and policy framework affect learning—including not just education and labor market policies, but also trade and industrial policies.

The first three parts of the chapter discuss *what is to be learned*, *the process of learning*, and *the determinants of learning*. We provide a taxonomy of the basic ingredients to a learning society. The next parts of the chapter look in depth at two of the major determinants of a learning society—spillovers and motivation. After then discussing some of the important impediments to creating a learning society, the final section of the chapter analyzes some of the key trade-offs in the design of a learning society.

Much of learning occurs within firms. The next chapter uses the ideas developed in this chapter to examine more closely what enables some firms to learn better than others and to explain why policies that help promote the industrial sector are likely to be particularly conducive to creating a dynamic learning society.

1. What Is to Be Learned

Most of this book focuses on learning how to use inputs better to produce outputs: how to increase productivity, getting more output per unit labor, capital, energy, or other resource inputs.

When we think of societal learning (and a society's learning capacities), we need to think broadly, not just in terms of particular production processes, upon which much of our discussion so far has focused. Societies also have to learn what products are best suited for their environment, an environment which is always changing, as both preferences and market conditions change. Sometimes, it is best to think of the outputs as certain "services" that are enjoyed by consumers, and innovation consists of new products that provide these services in better and less expensive ways.

Learning about Comparative Advantages

Some individuals in society are more capable than others, or at least better suited for particular jobs. Part of the role of the education system is to identify these absolute and comparative advantages. But some education systems perform these "learning" tasks better than others (see, e.g., Stiglitz 1975b).

Learning to Manage Organizations and Societies

One of the most important aspects of learning is learning to "organize," to manage collectivities of individuals. Large organizations can do things that small organizations can't, but managing large organizations requires knowledge that is different from that associated with managing small organizations. One of the main advances of the twentieth century was figuring out how to manage large research projects—divide tasks into components that could be undertaken by different groups, with pieces that could be subsequently put together.

In each arena of our society, there has been learning that has enabled our complex society to function. Keeping accounts is necessary for the functioning of a modern firm and a modern society. Thus, learning about accounting is essential; and keeping accounts for large organizations or in a modern large economy requires ever more complex systems.

Our complex society could not function without regulations, either. But again, there is considerable scope for learning—learning, for instance, how to regulate in ways which control externalities without imposing undue costs. The failure of banking regulations—both to ensure that financial markets perform the societal functions that they are supposed to perform and that they not impose the enormous adverse externalities that they did in the crisis of 2008—shows that there is still much to be learned.[3] This is part of a broader theme: learning how to make markets act like markets are supposed to behave.

This book argues that there are a number of government policies that could foster creating a learning economy and society. But managing these policies well may not be easy; there will have to be a learning process. Part of the learning process will be failures; decisions will be made which, at least in retrospect, look wrong or misguided. The conclusion that one should reach from these failures is not that the policy should be abandoned, but that one has to learn how to better manage the policy.

The crisis provided an illustration. It showed that monetary policies, as they had been formulated in the United States and many other countries, were badly flawed. Inflation targeting would not necessarily lead to macroeconomic stability. Financial markets were not necessarily very good at self-regulation. It was at least equally important that Central Banks focus on financial stability. (Indeed, it provided another lesson for developing countries: some of what they had learned from the developed countries as "good" policies and institutional frameworks was itself misguided. Some of the problems that they had experienced were not because they hadn't done a good enough job learning what was required to manage an economy well; it was because what their teachers had been teaching them was wrong.)

But the lesson to be learned from the failure of, say, America's monetary policy and institutions was not that monetary policy should be abandoned and the central bank shut down. Rather, it was that monetary policy and the central bank should be changed.

In particular, we argue here that there is an important role for industrial policies—policies designed to promote certain sectors and/or technologies. There have been failures of such policies. But there have also been notable successes. Countries will have to learn how to do a good job at conducting these policies, including what kinds of institutional arrangements are most conducive to success. But just as one can only learn how to better produce steel by producing, one can only learn how to conduct industrial policies by having such policies. There has to be learning by doing.

Firms and societies have to learn how to *compete*. They have to learn how to export as well.

In short, everything we do—as individuals, as organizations, as societies—requires learning. Things can be done better; we can be more successful in accomplishing our ends—in ways which require less resources and less time.

Learning Capacities and Learning to Learn[4]

Not only do firms (and societies) differ in their ability to transform inputs into output (i.e., they differ in their *knowledge*), they also differ in their ability to learn. Some individuals, firms, and countries are quicker at picking up changes that have occurred elsewhere, discovering knowledge that might be relevant, and adapting technology to their circumstances.

But just as *knowledge* itself is endogenous, so is the ability to *learn*. Some economic activities (conducted in certain ways) not only facilitate learning, they may facilitate *learning to learn*.

Paul MacCready's attempt to design a human-powered flight vehicle illustrates a recent instance of "learning to learn." He realized that key to designing such a vehicle was learning to learn. Previous attempts involved large investments based on often well-conceived theories, but when the vehicle crashed, there was no opportunity to make refinements. He focused on how to build a plane that could be rebuilt in hours. That enabled him to *learn*, to correct mistakes at reasonably low costs, and in short order, he was then able to construct the desired device.[5]

Learning abilities can, of course, be specific or general, and there may be trade-offs between the two: Some individuals may have an ability to learn quite generally, while others have developed more focused capacities. We can direct our efforts at enhancing specific learning abilities—that may serve an economy well if it is pursuing a narrow niche; or efforts can be directed at more general learning abilities, which may serve it well in periods of rapid transition and great uncertainty.

Learning for Development

Within developing countries, skills that are of especial relevance but particularly scarce are those associated with entrepreneurship.[6] One of the attributes of good entrepreneurs is their ability to learn and adapt. Some societies do a better job at learning who is better at this kind of learning, of selecting potential entrepreneurs.

So too, developing countries have to learn what products they are best capable of producing and are best suited to their conditions (see Hoff [1997] and Hausmann and Rodrik [2003]; later in the chapter, we discuss these issues further).

One of the key problems facing many developing countries is that they are exposed to high levels of volatility, and they have weak institutions for coping with this. Learning to manage risks is thus also important for successful development.

2. The Process of Learning

Some learning is a result of explicit allocation of resources to research and development, but much learning is a by-product of production and investment.

Learning by Doing

We learn by doing. We learn how to produce more efficiently by producing—and as we produce, we observe how we can do it more efficiently. There is ample empirical evidence supporting this hypothesis at the micro-level, both before and after Arrow's classic work.[7] Much of the formal analytics of this book is predicated on the assumption that much learning occurs by doing.[8] While this assumption greatly simplifies the analysis, it is a straightforward matter to extend the model, and at a few points, we show how this can be done.

How much we learn by doing is affected by how we do what we do. If we consciously experiment on the job, looking for alternative ways of doing what we are doing, we are likely to learn more than if we passively wait for a *eureka* moment, when we have a brilliant insight about an alternative way of doing what we have been doing.

Arrow's model itself linked learning with investment. Many advances in technology are embodied in new capital goods (Solow 1962b). The more "machines" that are made, the better the machines and the higher productivity. But interestingly, much of the cited evidence on learning by doing relates learning more directly to production, for instance, to the number of airplanes constructed.

But investments can promote learning and productivity improvements in other ways as well. Technological knowledge is embodied in machines, and a machine constructed for one purpose can often be adapted for quite another. It is not an accident that the Ohio Valley (stretching up to Michigan) gave rise to innovations in bicycles, airplanes, and cars. While the products were distinct, the development of these products shared some of the same technological know-how. This illustrates the principle that it may be difficult to identify ex ante what are "nearby" products, products such that advances in learning in one affect the other. (We will return to these themes later.)

New machines can also be a catalyst for learning. Computerization provides an important example. In the process of computerization, firms had to rethink their business operations, to codify much of what they had done without thinking. Through that process, they came to learn, to think about how much of what they did could be done better.

So can new production methods be a catalyst for learning. Just-in-time production not only served the function of reducing inventory costs. But following the motto "you only find out who is swimming naked when the tide goes out,"[9] just-in-time production exposed problems in the production process, forcing firms to address them. In a sense, it forced learning.

It follows from the fact that we learn by doing, that what we do and how we do it affects what we learn and the evolution of our economy and society. When a society focuses on learning how to save labor, reducing the labor input per unit output, it increases its capacities for this type of learning; conversely, if it chooses to focus on learning how to reduce the environmental footprint of products and production, its learning capacities in that direction can be enhanced.[10]

Learning by Learning

Just as we learn by doing, we learn to learn partly by learning. Hence, there may be a virtuous cycle: Countries that have managed to advance technology, providing more opportunities (and a greater necessity) for learning, may simultaneously enhance the ability to learn.

Learning from Others

We also learn from others, both in formal education and, even more importantly, in everyday contact with others. Knowledge is embodied in people and is transmitted by contact among people. This is especially relevant for what is called *tacit knowledge*, understandings that are hard to codify, to articulate as simple prescriptions that can easily be conveyed through textbooks or classroom learning. Workers move from firm to firm and thus convey some of the learning that has occurred in one firm to those in others.

Equally importantly, what we learn from others (or from achievements of others) can be a *catalyst* for our own learning: it may lead us to ask new questions or to see things in a slightly different way, and the result of this may be new insights, new learning.

Knowledge is also embodied in firms that supply inputs to multiple firms. What they learn in dealing with one firm in one industry may be relevant for another firm. There can be backward, forward, and horizontal linkages (Hirschman 1958).

The structure of the economy (including policies and regulations) affects the extent to which learning from others occurs. An economy with more mobility and openness is likely to be one in which there is more such learning. Some labor contracts are designed both to reduce mobility and to reduce the scope for ideas to be transmitted through mobility. Universities have traditionally been structured to maximize the extent to which there can be learning from others. More recently, however, the intrusion of IPR through the Bayh-Dole Act (which enables universities to appropriate some of the returns to research which goes on within them) may have provided increased incentives for secrecy and for less openness. On the other hand, the Internet has provided a technology which has facilitated learning from others.

More openness and mobility affects the flow of information through society, but it may have adverse effects on incentives to learn. This is another example of a kind of trade-off that is pervasive in the analysis of learning economies. In later chapters, we will show that the market solution—for instance, the level of mobility that results as part of a Nash equilibrium in which each firm tries to recruit workers from others but imposes restraints on others from recruiting from itself[11]—is not in general efficient.

One of the objectives of industrial policies (broadly construed, as described earlier in this book) is to facilitate learning from others. This is especially true in agriculture, particularly in developing countries, where "model farms" have sometimes been used to help disseminate best practices. One of the objectives of colleges and universities has traditionally been to facilitate learning from each other.

Learning by Trade

Trade, of course, facilitates interactions and thus learning. Advocates of free trade suggest that expanding trade is important because it facilitates learning. Successful exporters *have* to learn what it is that customers want; they have to learn about what competitors are supplying and figure out how to outperform. Domestic producers exposed to foreign competition from imports *have* to learn how to compete—how to produce products that are at least as good as those by foreign competition. More broadly, opening up to the rest of the world catalyzes learning and provides contacts from whom one can learn.

Later, we will explain why some trade restrictions may actually enhance societal learning. In making this argument, we are not arguing for autarky; we are not even denying that there are learning benefits from trade. What we are saying is that there are also learning benefits from domestic production and that the free trade literature has essentially ignored these benefits. We are also arguing that in assessing the learning benefits from trade, one has to be more precise in the analysis: (1) What are the sectors/products/technologies being traded? (2) What are the learning spillovers from abroad to the domestic agents? (3) What are the learning spillovers from those domestic agents to the rest of the economy? (4) What is the counterfactual learning: If the product had not been imported, or imported to the same extent, what is the level of learning that would have occurred? With what spillovers? Could the government have shaped the domestic production (e.g., the choice of technology) in a way which enhanced learning, more effectively than it can shape learning that is brought about through trade?

In the simple models presented in later chapters, we provide unambiguous answers to these questions, showing that there are contexts in which trade restrictions can help promote learning and hence raise long-run living standards to a far higher level than would have been achieved through free trade.

Technology and Learning Processes

Changes in technology affect what and how we learn (and what and how we *should* learn). One might caricature the "old" learning model as one in which the teacher pours knowledge, viewed as relevant at the time the kid is going to school, into the brain of the child, which he will then draw upon the rest of his life to solve problems. This model was never fully appropriate: at least the better schools prided themselves in also providing analytic skills critical for problem solving.

More recent years have seen a shift in focus to lifelong learning, with the recognition that what one is going to need to know in twenty years can't be adequately anticipated today. But the Internet and the vast store of knowledge instantaneously accessible there has changed matters further. Why store in the brain information that can be accessed in a moment? Some suggest that all that one needs to store in the brain is knowledge relevant to accessing information quickly, but that is clearly wrong. There is a plethora of information that flows over the Internet, and one must constantly make judgments about the quality (veracity) and relevance of the information. One must put the information received into context and be able to use it in conjunction with other information.

By the same token, much of what goes on in the workplace is constantly changing, and employers do not expect employees to come to them equipped to be a fully productive member of the workforce. The expectation is that there will have to be on-the-job training—and this on-the-job training will be ongoing. Hence, we can think of our lifetime education system as consisting of two parts, a formal part ("schools") and an informal part ("jobs and elsewhere"). The two are complementary as much as they are substitutes: if the first does its job well, it increases the returns to expenditures by the second. But unfortunately, there is little coordination between the two, at least in most countries, and hence formal schooling is often of limited relevance to on-the-job learning in many sectors.

What should be clear is that changes in production and knowledge technologies have altered the way we do and should learn, and a well-functioning learning society adapts itself to these changes.[12]

In the previous section, we discussed the learning associated with trade. But that too can be affected by technology. If a country produces a complex product (like an automobile), for which it imports most of the parts but makes some of the parts locally, there is *some* learning associated with the manufacturing of these local parts. But when the technology becomes highly complex, all of the components may be made and assembled abroad; even the technology for assessing whether the object is working effectively may be imported from abroad. There may be learning associated with how to use the technology but little learning relevant to production. The extent of spillovers may be lower, and thus the learning benefits of trade may be reduced.

3. The Determinants of Learning

A central thesis of this book is not just that learning provides the key explanation for the remarkable increases in standards of living in the past, but that the rate and direction of learning is endogenous, differs across countries and over time, and can be affected by the decisions of individuals, firms, and governments. The central policy issue of this book is how to enhance learning—how to create a learning economy and society.

With our previous discussion of the objectives and processes of learning as background, this section provides a taxonomy of the major determinants of learning: (1) learning capabilities; (2) access to knowledge; (3) the catalysts for learning; (4) creating a creative mindset—the right cognitive frames; (5) contacts—people with whom one interacts—which can catalyze learning, help create the right cognitive frame, and provide crucial inputs into the learning process; and (6) the context for learning.

Learning occurs at all levels within a society—individuals learn, but so do enterprises, and even governments. More generally, there may be "social learning," changes in societal beliefs that lead, in turn, through the political system, to different public actions.[13]

The analysis of these determinants provides the crucial ingredients for the design of a *learning architecture*: designing structures (e.g., firms, institutions, and frameworks in which they interact), policies, and societies more generally that promote learning and innovation. While some firms—the most innovative ones—have worried about how to design themselves in ways that promote learning within themselves,[14] surprisingly—given the importance ascribed to innovation in modern capitalism—the subject of how to design the overall economy to promote innovation has been given short shrift.[15] The objective of this book is to fill this lacuna.

Learning Capabilities

The most important determinant of individuals' learning is their capabilities, their ability to learn, and perhaps the most important determinant of that is education. As we emphasized earlier, individuals have to learn how to learn. Well-designed education systems (not those focusing on rote learning) are concerned precisely with learning to learn. As we have noted, modern education and labor market policies focus on "lifelong learning," enhancing the ability to adapt to an ever-changing marketplace. This facilitates individuals moving from one firm to another, with large private and social benefits to the ensuing flexibility. Since much, if not most, economically relevant learning occurs on the job, not in formal schooling, one should see formal education and on-the-job training as complements, with the former designed to enhance the productivity of the latter.

Much of traditional economics focuses on education's role in increasing human capital, the stock of knowledge embodied in individuals. It is typically measured by years of schooling. Our emphasis is quite different. Years spent on rote learning might (or might not) increase the stock of (even relevant) knowledge and, in that sense, increase productivity, at least temporarily, until that knowledge becomes obsolete. But such schooling would not necessarily increase the ability to learn—increasing capacities for lifelong learning—and could actually impede it, especially if, as part of such education, there is an attempt to inculcate ideas that are antithetical to science.[16]

Earlier, we emphasized trade-offs between static efficiency and dynamic gains. This is also true in education. In the short run, we might be able to impart more knowledge through an education system that demanded frequent testing of what students have learned and that focused less on enhancing analytic skills and cognitive abilities; but the learning capacities and creativity of those emerging from such an education system might be less.

AGE STRUCTURE Among the determinants of a society's (or firm's) ability to learn is its age structure—and possibly the age profile of its management structure. As the old saw has it, you can't teach an old dog new tricks. *On average*, younger individuals are more capable of (and open to) learning. In a sense, they have no choice: they have to learn the skills and knowledge that will enable them to succeed.

They also are likely to have greater incentives for learning. They are not invested in old ideas and ways of doing things; indeed, they even can have incentives to create new ways of doing things in which they excel.[17] They have a lifetime to benefit from such learning, and such learning gives them a competitive advantage over those who are older.[18]

One of the concerns facing Western societies (including Japan) is their changing demographics, in which the proportion of young people in the workforce will diminish markedly. The effects of this can be partially offset in economic systems in which there is a rapid pace of turnover of firms—with new firms, dominated by younger individuals, playing a vital role.[19]

Access to Knowledge

All knowledge builds on preexisting knowledge. As Isaac Newton described his own path-breaking work, "If I have seen further, it is by standing on the shoulders of giants."[20] That is especially true in our fast-moving innovation economy, where producing a complex product requires solving dozens, perhaps even hundreds, of problems. That is why access to knowledge is key to learning and to the further advancement of knowledge. We have noted the special importance of this in the process of development, given the recognition that what separates developing from developed countries is more a gap in knowledge even than a gap in resources.

Many aspects of the design of the economic system affect access to knowledge. The open-source movement is motivated by a commitment to access. As we noted, universities and modern science too work hard to maintain a culture of openness which ensures access to knowledge.

Later (chapter 12), we will discuss the ambivalent role played by intellectual property: while it may provide greater incentives for undertaking research, it simultaneously restricts access to and use of knowledge.[21]

Catalysts

Learning requires individuals and organizations to have the capabilities to learn, but individuals and organizations have to be spurred to learn. In a learning society, individuals are exposed to many *catalysts*. We use the term *catalyst* deliberately. As we have noted, much knowledge (innovation) builds on other innovations. But sometimes, one idea can incite new ideas—even if the new ideas do not "use" the old idea or build on it directly. In that sense, they are like a catalyst—a chemical that facilitates a reaction but is not itself used (or used up) in the process.

We learned from the discovery of rayon that it was possible to create (at affordable prices) synthetic fibers. Simply knowing this—even without knowing the precise way that rayon was constructed—can be an important spur to further learning and research. But we learn even more—we can be stimulated even more—from the disclosure of information that is contained in a patent application, even if, because of the patent, we cannot use (without paying for it) the patented product itself.

Advances in technology are among the most important catalysts to learning: We can learn more if there is more to learn. Policies, including government expenditures, that result in faster movements outward of leading-edge technologies mean that there is more for others to learn—if we do not impose impediments to their learning (e.g., through intellectual property).

Contacts

Earlier, we described the learning process, and central to the learning process is people learning from other people. These interactions both provide the knowledge input that is the basis of learning and provide the catalyst which enhances innovation.

Knowledge, in this sense, is like a (good) disease: it can spread upon contact. But some kinds of contact are more likely to lead to the transmission of knowledge than others. Some of the people who might possibly come into contact with the knowledge are "susceptible"; i.e., they are more likely to learn, to use the knowledge, and perhaps even to develop it further. And some types of economic structures can facilitate individuals coming into contact with each other—while other structures can impede contact. Universities and research institutes try to create an interactive environment to enhance the range and depth of contacts. Economic systems that encourage mobility may increase the extent of contacts that bring with them learning and catalyze learning. Structured interaction can be even better—that is, organizational architectures that help bring individuals who might stimulate each other into contact. Some firms have policies of regularly shifting their employees, partly because in doing so, it facilitates the spread of ideas across the firm and promotes learning within the firm.

Traditionally interactions were largely affected by geographical proximity, and this helps explain the development of learning clusters—locales at which learning, especially in particular areas, occurs at a more rapid pace than elsewhere. The strength of these local interactions provides one of the main sources of agglomeration economies—why it makes sense for certain activities to congregate together in particular places. The "localization" of contacts and the ability to communicate provides one of the explanations for why knowledge moves more freely within a country than across national boundaries. Differences in language are a barrier to the movement of ideas; sharing a common education system can, by contrast, facilitate the movement of ideas.[22]

One of the benefits of globalization is that it has expanded individuals' exposure to new ideas.[23, 24] Contacts, of course, don't have to be face-to-face; the Internet has vastly expanded the ability of individuals to communicate with each other, improving both access to knowledge and the possible range of contacts.[25]

Cognitive Frames

Individuals and firms have to adopt a cognitive frame, a mindset, that is conducive to learning. That entails the belief that change is possible and important—and can be shaped and advanced by deliberate activities.[26] Part of the reason for the relative stagnation in living standards that persisted for thousands of years before the industrial and learning revolutions was that there was not this cognitive frame.[27] And of course, the absence (or slow pace) of change meant that these beliefs were self-reinforcing. In a world with little change, there are few catalysts to learning, and little effort is spent on creating change and adapting to it. And because people were so bad at adapting to change, because they had not learned how to learn and did not have the institutional structures that helped them bear the costs of change, there was often large resistance to change (e.g., the Luddite movement of the nineteenth century).[28]

In the West, the Enlightenment—with its belief in science and rationality, careful experimentation, and close inference—was pivotal in creating the learning mindset. It represented a marked departure from a mindset that saw truth as being revealed from on high. It is curious that while those in developing countries are striving to embrace the scientific method, in the United States, the country which has led the world in the development of technology, large swaths of the population cast aspersions on the results of modern science, most notably evolutionary theory and climate change.[29] Policymakers trying to promote learning and the advancement of science and technology often seem to face the task of relitigating the Enlightenment.

How a learning mindset is created is a complex matter.[30] Fundamental beliefs, such as those that we have been discussing (such as that change is desirable and can be created), are, to a large extent, social constructions. We believe what we believe partly because those that we talk to believe similarly. This can result in societal rigidities—it is hard for any individual to change his mindset on his own, or for any single individual to bring about a change in the collective mindset. But such changes are essential for development—for the transformation of countries from stagnation to growth, to becoming a learning economy.

At the same time, beliefs have to confront reality. A large gap between beliefs and reality provides a strong impetus for a change in beliefs. But ideas matter and have a life of their own. The spread of the Enlightenment and the scientific method was partly based on their ability to provide convincing interpretations of observations that seemed otherwise inexplicable, to make predictions that could not have otherwise been made, but it was partly based, too, on the power of the ideas themselves.[31]

While we may not fully understand the spread of the "learning mindset," what is clear is that education is pivotal. A well-designed education system can help create the right cognitive frame, but there are also education systems that can "inoculate" individuals against the Enlightenment and the development of a learning mentality.

Context: General Observations on a Learning Environment

Learning occurs in a context. Most learning occurs within a firm. Some environments (e.g., the culture of some firms or societies) are more conducive to learning than others. They can help create learning capabilities and a learning mindset and establish networks of contacts that provide strong catalysts for learning. Some environments can stifle learning, failing to develop learning capabilities, inhibiting the flow of knowledge, and making it difficult to put to use learning that occurs, and because much learning, as we noted, is a result of "doing," this inhibits further learning.

The extent of learning can be affected both by the macro-economy and by the structure of the firms in which individuals work. Later in this chapter, we'll discuss two key aspects of the learning environment, spillovers and motivations, and we'll discuss in greater detail some of the other more salient aspects of the learning environment in the next chapter. (Later chapters will discuss still other aspects of a learning environment, touching on social protection, labor legislation, and laws affecting finance and investment.) We note that both the macro-environment and the system of social protection—as well as the level of inequality and other attributes of the economic system—can affect learning. For instance, individuals who are preoccupied with survival or who face high levels of stress typically cannot learn as well as those who have a modicum of security. The prevalence of stress and anxiety is likely to be greater in societies in which there is a low level of trust.[32]

Here, we note that there may be multiple societal equilibria. More dynamic societies with larger change create a greater demand for learning; they reward those with learning capabilities more and incentivize individuals to acquire those skills and mindsets. Societies with little change put little value in these skills and thus fail to incentivize individuals to acquire them. The result is that there is little change. (Assume, for example, that some individuals are better in bureaucratic skills, others at innovation. There is one societal equilibrium which can be thought of as "bureaucratic," where those who are better in bureaucratic skills thrive, and individuals within this society learn to better manage bureaucratic processes. Other societies become more innovative. In such societies, those with more innovative abilities prosper, relative to those who are better at managing bureaucratic processes, and individuals develop the learning capacities that enable them to prosper in such societies.)[33]

4. A Closer Look at Learning Spillovers

We emphasized earlier that there are important positive externalities from learning. Such spillovers are pervasive and large, and they are larger in some industries than in others. Obviously, markets will not take into account these externalities—a critical market failure which is at the heart of the analysis of this book.

Localized Learning

There are many aspects of learning spillovers. As Atkinson and Stiglitz (1969) noted, learning is localized: it affects production processes that are similar to those for which there has been learning more than it affects production processes that are markedly different. Improvements in a capital-intensive process of making textiles may have little impact on hand-loom technology. But the learning is not limited to a single process and related processes for a particular product. Innovations in one sector may benefit seemingly unrelated sectors, because the production of any good involves many stages, and some of the stages may involve processes that are similar to those used in another seemingly distinct sector. Sectors that are, in one way or another, more similar may, of course, benefit more. Indeed, the spillovers may be greater to other products using analogous technologies than to firms using dissimilar technologies within the same industry.

The spillovers involve more than learning about technology. There are especially important spillovers in methods of production. Inventory control and cash management techniques affect virtually every firm in an economy. Just-in-time production or assembly lines are examples of production processes that affect many industries.[34]

There are, by the same token, "institutional" spillovers. The development of a financial sector suited to serve, say, the manufacturing industry may have enormous benefits to other sectors of the economy. Many of these spillovers can be economy-wide. Similarly, improvements in the education system, necessary for an effective industrial sector, too can have benefits for the service sector or the agricultural sector; indeed, the benefits are likely to be economy-wide.

The spillovers involve not just technologies but people. Improvements in skills in one sector have spillover benefits to other sectors in which analogous skills are employed.

The theory of localized learning suggests that spillovers may flow more naturally not only from one technology to other similar technologies but from one product to certain other products. This may due to institutional factors—people flow more frequently from firms producing one product to those producing another—or due to technological factors, related to the similarities in certain aspects of the production processes. Hidalgo and colleagues (2007) characterize the product space, attempting to identify where there are the most significant spillovers, e.g., which sectors entail similar “capabilities.” Presumably, if two products entail similar capabilities, learning that enhances a particular capability in one sector will have spillover benefits to related sectors for which that same capability is relevant.[35] Capabilities, in this sense, can include not only worker skills but organizational learning and institutional developments. They map out industries that seem related, i.e., where the development of one industry is associated with the development of another, but this describes the spillovers that exist under current institutional arrangements. It does not describe the spillovers that might occur under alternative institutional arrangements, e.g., with a more active industrial policy.

GEOGRAPHICAL AND CULTURAL LOCALIZATION[36] There are other aspects of localization: geographic and cultural. Learning is geographically localized, in part, because knowledge relevant to one locale is less so to another. It is also localized because information flows are localized. When people are spread apart geographically, there is a less dense set of contacts, and because of language and cultural differences, communication may be less effective. Transportation networks, too, facilitate interactions, say, within a country rather than across national boundaries. Because learning itself is a complex process, one may be able to learn more easily from those who speak the same language and are attuned to how one thinks and perceives the world—and so there may be larger learning spillovers.

Thus, learning spillovers may be larger to countries (regions, locales) that are *similar* in some fundamental ways. This is obviously the case in agriculture, where agriculture improvements suitable for one locale may not be for another, where soil types and rainfall and other aspects of climate may be markedly different.

Geographical localization is one of the reasons that knowledge flows less freely across borders than it does within borders—something which should be apparent from the large disparities in productivities across countries and which is central to the analysis of later chapters.

While geographical localization means that some learning that is relevant in one locale will be *less* relevant in another, most changes in technology and many changes in institutional learning, however, *could* confer benefits across borders. The extent to which that is the case may depend on the level of skills (human capital) and the institutional arrangements.

In chapter 4, we will note another aspect of "localization," the fact that knowledge or learning flows more easily within a firm than across firm boundaries. Multinational firms have facilitated the movement of knowledge across national boundaries—but often have been effective in restricting the flow of knowledge within firm boundaries, limiting, as a result, the benefits of foreign direct investment (FDI). Chapter 11 discusses government policies which can enhance the learning benefits that can be reaped from FDI.

LOCALIZED LEARNING TO LEARN Not only is knowledge (learning) localized, so too may be learning how to learn. Countries (firms) are likely to learn about this learning more easily from neighbors. It is not an accident, in this view, that Japan's neighbors were the first to learn how to learn—they observed, and imitated, what Japan had done. This knowledge of learning how to learn then diffused around Asia. Korea was attentive to the policies that Japan had pursued that had allowed it to close the knowledge gap between it and the more advanced industrial countries. (Some of these policies will be the subject of subsequent discussion.)

There is another aspect of localized learning to learn. Learning involves specialized skills, and improvements in learning in one area may be at the expense of learning to learn in other areas. Western firms have learned how to save labor—even when there is a high unemployment rate, so there is a high social cost to such labor-saving innovation. But they have not done well in learning how to protect the environment and reduce their resource footprint.

LOCALIZED LEARNING, CONVENTIONAL ECONOMIC THEORY, AND WHY HISTORY MATTERS The fact that much learning occurs within a firm (or a country) and does not easily or costlessly move across (firm or country) boundaries has profound implications for conventional economic theory. It means that firms (countries) necessarily have different production functions. To assume that they have the same knowledge is as silly as assuming that they have the same factor endowments. Indeed, much of the modern theory of trade is predicated on the imperfect mobility of factors of production but the perfect mobility of knowledge (that is, it is *assumed* that production functions are the same); but the movement of knowledge across boundaries (whether of firms or countries) is even more difficult than the movement of factors. Indeed, it is imperfections of information that represent the most important impediment to the movement of capital: if information flowed perfectly, it is arguable that so too would capital.

If, of course, at least some learning/knowledge never moved, then we could still approach the problem of resource allocation using conventional frameworks. Given the state of knowledge of each enterprise, we would allocate factors to ensure marginal returns are the same for each use. But a central message of this book is that that approach is misguided: we can affect the flow of knowledge; we can affect learning; these are affected by economic policies, institutions, the design of economic structures, and resource allocations (both sectoral allocations and choices of technologies).

The fact that learning is specific to a particular technology (a particular technique) means that, in a sense, the set of techniques available at any point of time is likely to consist of techniques presently used or used in the past. Of course, there are a range of techniques which, with additional investments, could be developed. But it means that, for practical purposes, the distinction between movements along an isoquant and "technical" change may be less important than is usually assumed.

It means too that history matters—in a way that is not the case in the standard model. Figure 3.1. shows an isoquant of a firm, entailing two techniques of production, a capital-intensive technique labeled *A* and a labor-intensive technique labeled *B*. The notion of localized technical change means that an improvement in *A* may leave *B* unaffected, and vice versa. Thus, common formulations of technological change, which see it as shifting the production function in some smooth way (reducing the input requirements for *all* technologies), are fundamentally misguided.

There are other elements of strategy: Though one can learn more easily "locally," i.e., about products and processes that are similar to those that one currently employs, there is *less to learn*. Thus, we spoke earlier of the large knowledge gap between developing countries and developed. They have more to learn. *If* they can learn how to learn about these more advanced technologies, because there is more to learn, there may be more learning—higher rates of productivity growth. At a given set of learning capabilities, the relationship between productivity increases may be complex, e.g., exhibiting an initial phase of increasing returns with respect to the size of the gap (one can learn very little if

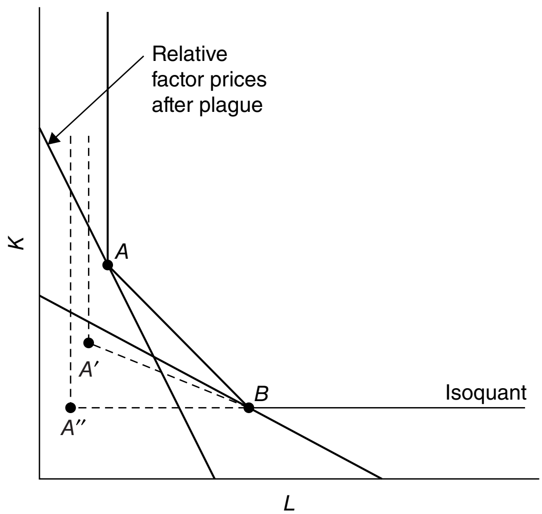

Figure 3.1 Localized technical progress: History matters

Improvements in the capital-intensive technology leave the other technology unaffected. Eventually, the capital-intensive technology dominates the other technology. (Assume that initially, the economy is in a steady-state growth path in which it uses *B*, with capital expanding at the rate of the labor force. Assume, further, that this technology is one which gives rise to little learning—per capita income stagnates. Then the country faces a plague, which wipes out a large fraction of the labor force. Given the increased labor scarcity, wages rise and the economy shifts to the capital-intensive technology *A*. But there is considerable "learning by doing" associated with *A*. As *A* is used, it improves, moving to *A′* and *A″*.)

there is virtually no gap), a phase of diminishing returns and, finally, if the gap is too large, even decreased learning.

The fact that there is more to learn from more distant technologies but weaker learning capabilities poses a difficult strategic choice: Should one follow a strategy of incrementalism (small changes) or of a "big leap"?

The appendix to chapter 2 that was included in the unabridged version of this book described "equilibrium knowledge gaps." There are large costs associated with overcoming knowledge gaps, and if the knowledge gap is too large, it may not be optimal to do so; that is, it pays a firm (country) to remain a laggard, absorbing knowledge (with a lag) as it filters down from those in technological leadership.

There is one further factor that may be relevant to a big leap: conventions, institutional arrangements, and mindsets are forced to change, thus facilitating the process of change. Rigidities in these represent a major impediment to learning and are the focus of the next section of this chapter.

5. Impediments to Learning

Earlier, we referred to the importance of the Enlightenment in helping create a learning society. But even within an enlightened society, there are barriers to learning at the individual, organizational, and societal level. Understanding these barriers—and doing what can be done to remove them—is central to creating a learning society.[37]

Biased Perceptions

Learning involves seeing what works and doesn't work—learning from experience. In science, we learn through controlled experiments. In real-life situations, it is often difficult if not impossible for individuals to conduct controlled experiments. We attempt to make inferences from our experiences about what works and what doesn't. But the inferences we make are themselves affected by our beliefs. What information we process, the way we process it, and the weight we give to various observations are all affected by our prior beliefs. The tendency to see the world through lenses that are shaped by our prior beliefs—to discount information which is inconsistent with those beliefs and to see information that is consistent with those beliefs as particularly salient—is called confirmatory bias (see Hoff and Stiglitz [2010] and the references cited there). The result is that we can live in a world which Hoff and Stiglitz describe as an "equilibrium fiction." The world—as we see it—confirms our prior conceptions.

The notion that a belief system can be a fiction—even an equilibrium fiction—does not mean that it is necessarily bad. Long ago, Frank Knight (1921) argued that entrepreneurs systematically overestimate returns to innovation, or (to use more modern terminology) exhibit irrational exuberance. This irrational exuberance can be a major spur to innovative activity, offsetting in part the underinvestment that results from firms paying too little attention to the externalities to which their innovative activity gives rise. The irrational exuberance of the dot com bubble left in its wake a host of thriving enterprises (such as Google) that might not have otherwise obtained financing. The excessive investment in fiber optics played a key role in spurring India's technology boom, as the cost of interconnectivity fell dramatically (see Stiglitz 2006a).

At the same time, shorter term bursts of irrationality—fictions that come to be believed and are seemingly validated by the selective framing of evidence[38]—play a key role in the booms and busts (both in credit and equity markets) that have marked capitalism from the beginning.[39]

Belief Systems as Social Constructs

Belief systems themselves are to a large extent social constructs: What each individual believes is affected by what others believe. It is not just belief systems that are social constructs. The prisms through which we see the world are largely socially determined. (This is sometimes referred to as *preconfirmatory bias*.) Sociologists and anthropologists (such as Mary Douglas [1986]) have long argued for the need to incorporate belief systems into our understanding of how societies (including economies) perform. Firms are, of course, mini-societies, and understanding their behavior requires an understanding of the belief systems of the firm, a key ingredient in what is commonly called corporate culture.

Countries don't exist in isolation, and not everyone within a country shares the same belief systems. Belief systems within a country are affected by those outside and especially so by those that are nearby and similar and with whom the members of the country are closely linked. So too for a firm. And if the linkages of members of a firm with members of other firms is more dense, the belief systems of outsiders have a greater impact; it is harder for those within a firm to have a belief system that is markedly at variance with those outside. With globalization, those in one country encounter, and have to come to terms with, other belief systems.

Cognitive Frames and Learning

Earlier, we discussed cognitive frames as a critical determinant of learning. Our cognitive frames affect *whether* we learn. There is a "learning" cognitive frame, and that enhances our ability to learn. In some societies (in some firms), there is an attempt to constantly assess whether what is being observed is consistent with prior beliefs and models, and when it is not, to change beliefs and models. Other societies (firms) are far more conservative. More weight is given to inherited truths, and such societies resist evidence that contradicts those truths. (Some of our earlier discussion provides a partial explanation of the differences; societies with more educated individuals have a greater capacity to learn and are more likely to learn.)[40]

Because what one individual believes is affected by what others believe, it is hard for belief systems to change. No individual or firm on its own can change societal cognitive frames; the result is that cognitive frames are a major source of institutional and societal rigidities. If a society is trapped into a belief system in which particular categories of individuals (women, individuals ascribed as lower caste or assigned to a particular ethnic group) are viewed as less productive, they will be treated in that way, there will be less investment in, say, their education, and their behavior may reflect that treatment (thus partially validating the differential treatment). Even if there were some isolated individuals who understood the notions we have just presented, including that of an equilibrium fiction and what gives rise to it, it would be hard for them to move the society to a different (say, nondiscriminatory or prolearning) equilibrium. So too for a firm: Though the head of a firm may have disproportionate influence, what she can do is circumscribed by the beliefs of others within the firm, who almost inevitably are closely linked to those outside the firm, and the (head of the) firm will have at most a limited ability to change their beliefs.

These cognitive impediments to learning and change are reinforced by economic interests: Change always has winners and losers, and losers have an incentive to mount a challenge to the change. They have an incentive to see the world through a lens which sees the change in a less positive light.[41]

Third, cognitive frames affect not only the extent of our learning but *what* we learn. As we argued before, those with a strong ideological commitment to a particular view will resist information that seems to contradict it—they will either totally ignore it or, if they can't ignore it, discount it.[42] The world is a complex place, and we can almost always put a spin on what we observe to make it consistent with prior beliefs—and this means that we can, at least for a long while, not learn from what is occurring.

In spite of these impediments to change and learning, change and learning do occur. The dynamics of change are themselves complex, affected by changes in the outside world (including changes in technology and advances in science) and by economic interests. But unlike naïve Marxians, we do not believe that economic interests alone drive change. Change is often affected by the evolution of ideas, and particularly of overarching beliefs.[43] Once the Enlightenment notion that "all men are created equal" was accepted (however that idea came to be accepted, whatever the drivers), it was no surprise that it evolved in directions that brought within its ambit women and slaves.

The uber-ideology of the Enlightenment—the questioning of authority and the belief in meritocracy, the notion that change is possible and desirable, the respect extended to science and technology—have created preconditions that are favorable to the creation of a learning society and to learning institutions (firms) within our society.

Finally, not only is it possible for belief systems to change, belief systems are malleable—though not perfectly, and not instantaneously. Thus, over time *we can help create a learning society*. We can do this partly by understanding the limitations in our perceptions—how our perceptions and learning are shaped (e.g., by ideologies, by confirmatory and preconfirmatory biases). We can do it partly by creating a learning culture—a culture that respects and reflects science and the values of the Enlightenment, including the questioning of authority. We can do it partly by rewarding successful learning. Both the public policies described in this chapter and the policies pursued by individual firms can, if appropriately designed, assist in creating a learning society. By the same token, the *wrong* policies can impede the creation of a learning society.

Impediments to the Transmission of Knowledge

There are other impediments, both natural and "manmade," to the creation of a learning society, most notably with respect to the transmission of knowledge. Knowledge does not diffuse on its own; it has to be transmitted and received, and there are barriers at both ends.

The previous section described some of the impediments to learning, to extracting information out of the cacophony of signals with which the individual is bombarded. There are also barriers on the "sending" side. It is plausible that a market economy engages in excessive secrecy (relative to the social optimum). This, of course, has been the contention of the open-source movement. Collaborative research in the open-source movement is still economically viable, both because there are still economic returns (e.g., because of the tacit knowledge that is created by the learning/innovation process itself and from the advantages that come from being first in the market) and because there are important noneconomic returns to and incentives for innovation, which we will discuss briefly.

But these impediments are augmented by legal frameworks (including prosecutorial norms and decisions) that have been increasingly adopted. For instance, intellectual property regimes can not only create an impediment to the transmission of knowledge but encourage a culture of secrecy and a lack of openness. And while the disclosure requirements of patent laws have as their intent the dissemination of information upon which further innovation and learning can occur, in practice these requirements have been ineffectively enforced, and some firms (e.g., in software) actively promote weakening them. We will discuss these issues further in chapter 12.

It is worth noting that the "ideology" that emphasizes the comparative efficiency of the private sector in all matters, including research, and that stresses the importance of monetary incentives in promoting private sector activities, has in large measure led to these policies, the effect of which may in fact be counterproductive. This is itself an example of what may be an equilibrium fiction: Those who believe this believe that the "evidence" supports that belief. To them, any observed deficiencies in the performance of the economy must accordingly be traced to some government intervention. It was not the market's irrationality that led to the housing and credit bubble, but rather government's encouragement of home ownership among the poor. If the private sector appears less innovative than it should (or less innovative in one place than in another), it is because of government regulations that stifle innovation.[44]

The proclivity for secrecy by private firms may be only one of the reasons that the state has played a central role in creating an innovative economy and a learning society (see Mazzucato 2013). Innovation is very risky and often entails large investments—even large firms are risk averse; capital markets are imperfect—especially when it comes to investments like R & D that are risky and can't be collateralized; and the most important innovations have large societal spillovers. Hence, as we explain in chapters 5 and 6, there is a strong presumption that there will be underinvestment in research, and especially in the kind of basic research from which all else flows.

6. Motivating Learning

Because learning requires effort and resources, individuals and firms have to be incentivized to do research and to learn. But it is important to realize that it is not just monetary (pecuniary) rewards that are important. Indeed, much of the most important advances are motivated by curiosity, or by a desire for recognition by one's peers (see David 2004a, 2004b; Dasgupta and David 1994), or by the excitement of solving a difficult problem that no one else may have even posed, let alone solved. (In fact, it should be apparent that the first sentence of this paragraph is itself a reflection of a prevailing way of framing the issue by economists: Many individuals, particularly those who are successful scientists, *do not* necessarily have to be motivated to learn. The pleasure of learning is its own reward.)

There is a large literature suggesting that such intrinsic rewards can be a far stronger motivator than extrinsic rewards, like money (see, e.g., Stiglitz [2001a, 2012b] and the references cited there). Most of the most important discoveries (such as the decoding of DNA) were motivated not so much by financial rewards as by these other factors.[45] And to the extent that extrinsic rewards do play a role, it is more from peer recognition than from monetary returns.

Appropriability

For those who believe that the major issue in motivating learning is financial rewards, the question of the appropriability of returns becomes central: Only a fraction of the social returns to innovation can be captured by the innovator. There are learning spillovers, externalities. This suggests that, to the extent that learning depends on financial returns, there will be underinvestment in learning. (Later, we will explain that there are many circumstances in which the returns to innovation may well exceed the social returns; much innovation in a market economy can be thought of as rent seeking, and some of the returns to innovation represent rents that otherwise would have accrued to others. This implies that there may be too many resources allocated to such activities.)

Spillovers occur even in the presence of a patent system. Many advances cannot be patented (many advances in mathematics, for example), and the benefits of much of what is learned in the process of research cannot be appropriated. Indeed, the disclosure requirements of a patent are intended to enhance these societal benefits.

An idea like just-in-time production, replaceable parts, or assembly lines spreads quickly throughout the economy and can't be protected by intellectual property laws. Firms may engage in experiments, e.g., about what products will be well-received by consumers, but successful experiments may quickly be imitated, so the benefits of such learning may not be appropriated by those engaging in the experiment.[46]

Consider an "experiment" to discover whether conditions in a country are particularly suitable for growing a particular kind of coffee. If the experiment fails, those who conduct the experiment lose money. Because learning what grows well in a particular climate with a particular soil is information that is not patentable, if the experiment succeeds, there may be quick entry. The country benefits, but the "innovator" can't capture much of the returns. As a result, there will be underinvestment in this kind of experimentation.

A similar argument holds for why private markets will lend too little to new entrepreneurs. There is a learning process in the discovery of who is a good entrepreneur. A bank that lends money to a young entrepreneur who proves her mettle may find that the entrepreneur can easily be poached away by a competitive bank. Assume in the initial period that it cannot be ascertained who is a good entrepreneur and will repay her loan and who is a bad one and will not. The bank loses money on the bad entrepreneurs, but may not be able to be adequately compensated by "excess" returns from the good entrepreneurs. Because of the threat of poaching, the interest rate that the bank can charge a good entrepreneur (after the entrepreneur has demonstrated her success) will be limited to the competitive rate. But Stiglitz-Weiss adverse selection and adverse incentive effects limit the interest rate that can be charged in the initial period, which implies that there will be limited lending to new entrepreneurs.[47]

Market responses to the appropriability problem can impede learning—it leads to more extensive patenting and greater secrecy. As we explain later, the former can result in a patent thicket, imposing significant barriers to the use of knowledge. The latter can lead to impediments in the transmission of knowledge and undermine the open architecture that has traditionally been so important in the advancement of knowledge.

7. Trade-Offs

Previous sections have described a number of factors that can contribute to creating a learning society. But there are a number of subtle and complicated trade-offs.

Trade-Offs Between the Efficient Utilization of Knowledge and Incentives to Produce Knowledge

One trade-off centers around the benefits arising from the free flow of knowledge and the attenuation of (financial) incentives to learn: With the free flow of knowledge, it becomes more difficult for someone investing in knowledge to appropriate the returns. The issues are the same as those at the root of the Grossman-Stiglitz (1976, 1980) critique of the efficient markets hypothesis: If knowledge were perfectly transmitted, there would be no incentive to expend resources on gathering and producing knowledge. There would be underinvestment in knowledge creation (and in the case of developing countries, gathering knowledge from others). One of the costs of relying on the private finance of knowledge production is that it *must* entail the imperfect transmission of knowledge. And indeed, this is one of the advantages of public support for the creation of knowledge.

There can then emerge a number of complex and subtle trade-offs, at the level of the individual, the firm, and the economy: Not sharing knowledge will typically mean not receiving knowledge, or at least not as much knowledge. Knowledge is often "traded" not through market mechanisms, but in a process better described as gift exchange.[48] In an academic community—out of which the most important advances occur—there are tacit understandings about the culture of exchange. An individual who did not share would be cut off from her colleagues, and the likelihood that (at least in most fields) she could make a significant breakthrough would be limited, since every innovation is based on dozens, or hundreds, of smaller ideas and concepts.

An organization might want to keep knowledge that it produces "private" so that it could appropriate more of the returns, but it wants knowledge to be shared fully within the organization, so that anyone in the organization can build on it. But here too there are trade-offs: The more individuals in the organization that have the knowledge, the more likely that the information will leak out.

We noted earlier that the free mobility of people, ideas, and products helps disseminate ideas and can be a catalyst for learning. But again, such mobility may make it more difficult to appropriate the returns to investments in innovation. That is why many firms impose mobility restrictions in contracts with employees: It may lead to short-run inefficiencies in the allocation of labor, but these may be partially offset by long-run dynamic benefits—or at least firms insisting on these contracts believe it enhances their long-run profits.[49] (The appendix to chapter 4, which was included in the unabridged version of this book, explained why it is likely that the market equilibrium may result in insufficient mobility.)

Intertemporal Trade-Offs: Static Inefficiencies Versus Dynamic Gains

The earlier discussion of the process of learning highlights the importance of intertemporal trade-offs in the process of learning. Learning is an investment. We normally have to sacrifice current consumption—the additional experimentation from which we may be able to learn more comes with a cost. Even if (as in the simpler models presented later in this book) learning followed automatically from production, we can learn more—lower future production costs—by producing more today than we would have produced from the perspective of short-run static efficiency.

Investments in "technology capital" or "knowledge capital" are, in this respect, much like investments in human and physical capital, with one important difference, to which we have already alluded: Because of the importance of knowledge spillovers, there is a presumption that there will be underinvestment in learning (in technology capital), while there is no such presumption concerning human and physical capital. More precisely, as we will explain in later chapters, social and private returns are not likely to be well aligned.

As in the case of the savings rate, there is thus (often) a short-run versus long-run trade-off. But in the case of the savings decision, there is no general theoretical presumption that government intervention is desirable—future generations are normally expected to be better off, so asking the current generation to make still further sacrifices so that future generations could be still better off is problematic. Moreover, even with such sacrifices, the increase in the growth rate is only temporary. In the case of learning, there is a presumption (shown more fully in chapter 6) that the market allocation is inefficient and that government intervention can enhance societal well-being; and intervention can lead to a permanently higher growth rate. *It may be desirable to distort resource allocations in the short run, to operate below the production possibilities curve, or to force consumption patterns that do not maximize short-run "utility," in order to achieve dynamic benefits—higher growth rates*. This trade-off is at the heart of the analysis of this book.

Indeed, if there were not such a trade-off, one could view growth and efficiency policies as complementary. The neoclassical (Washington Consensus) policies move the economy to the production possibilities curve as quickly as possible, and then the growth policies move the production possibilities curve out as rapidly as possible. If the first policy is successful, then the benefits of the second policy—say, an increase in the growth rate of potential output (the rate at which the production possibilities curve moves out) by a given delta (denoted by Δg)—are even greater, since the level of output which is multiplied by Δg (the change in growth rate) is greater.

Much of modern growth policy analysis is based on this kind of dichotomy. Growth theory is based on the "supply" side, assuming that the economy is on its production possibilities curve—i.e., resources are fully used and allocated efficiently. This book will explain why that approach is misleading, if not just wrong.

Our "learning society" perspective even sees the reason that the economy might perform below its production possibilities schedule differently,[50] and accordingly, it sees the short-run versus long-run trade-offs differently. As chapter 2 explained, even well-functioning economies operate well within the output limits set by a traditionally defined production possibility frontier. The gap arises not so much from market distortions (associated, say, with taxes) but from disparities in knowledge and practice that exist even *within a country*. For most firms, as we noted in the last chapter, there is a significant gap between their productivity and "best practices." Current output can be expanded, therefore, without significant creation of new technology through the better use and wider deployment of existing technologies. If those who use less-efficient technologies can only "learn" best practices, there will be large increases in output. Output can be expanded by the more-effective deployment of labor within a firm, if only firms learn how better to use their workers—not just the reallocation of labor across firms caused by labor market rigidities. Typically small detailed changes in production processes can lead to large increases in productivity, without significant improvements in the quality of labor emerging from formal school systems. To assume that knowledge flows freely within and across firms (or across national boundaries) is as unrealistic as to assume that production could occur without inputs. Knowledge is as important a constraint on output as conventional inputs.

But policies that focus on static inefficiencies and their removal may, at the same time, reduce incentives for learning, or even impede learning. There can be, as a result, a trade-off between static inefficiency and dynamic efficiency. For instance, as we have repeatedly noted, intellectual property regimes (or other impediments to the free flow of knowledge) mean that knowledge is not being efficiently used (see chapter 12). Such impediments contribute to static inefficiency. But they *may* lead to enhanced incentives to research—to dynamic efficiency. But as chapter 12 also explains, poorly designed intellectual property regimes can be a lose-lose situation—less efficient use of knowledge today, less innovation and less growth in the future. Earlier in this chapter we noted another example of such a trade-off—creating a less competitive banking sector may introduce a static inefficiency, but at the same time, it reduces the chance of a good entrepreneur being poached away, and hence may lead to more lending to new entrepreneurs, enhancing the dynamic efficiency of the economy.

In each of these instances, of course, there may be policies which change the nature of the trade-off, that is, induce as much learning with a lower loss of static efficiency.

There are many other trade-offs in the construction of a learning society. Learning abilities can, for instance, be specific or general, and there may be trade-offs between the two: We can direct our efforts at enhancing specific learning abilities, which may serve an economy well if it is pursuing a narrow niche; or efforts can be directed at more general learning abilities, which may serve it well in periods of rapid transition and great uncertainty.[51]

8. Concluding Comments

This chapter has focused on several of the critical determinants of learning, in the hope that by understanding better the factors that affect learning, we can be more successful in creating a learning society—in the case of a developing country, one which more quickly closes the gap between itself and the more advanced countries; in the case of a developed country, one which moves out the frontier of knowledge at a faster pace; in the case of all economies, one which reduces the gap between average and best practices. It should be obvious that the conditions (the institutions, the legal frameworks, the contractual arrangements, etc.) which facilitate learning for different economies facing different circumstances may differ markedly. In particular, those which are appropriate for an economy trying to catch up may differ markedly from those appropriate for one at the frontier—trying to move that frontier out further. (Later chapters will illustrate this.)

There are several questions to which we have so far given short shrift. Much of the learning in our society occurs within firms. What are the determinants of the learning that goes on within a firm? Are firms in some sectors better at learning? Is there a tendency for firms in some sectors to provide greater externalities to the rest of the economy? Are there macroeconomic conditions that facilitate learning? And what policies can help bring about those macroeconomic conditions? We turn to these questions in the next chapter.

CHAPTER FOUR

Creating a Learning Firm and a Learning Environment

THE PREVIOUS chapter described many of the ingredients of a "learning society": what is to be learned, how learning occurs, and the central determinants of learning, including the role played by learning spillovers and appropriability. Among our central objectives was to describe some of the components of an economic architecture (including all the attendant institutions and laws, such as intellectual property rights) which would best facilitate learning.

This chapter focuses on creating two critical aspects of this learning architecture: a learning firm and a learning macro-environment.

1. The Learning Firm

A subproblem within the systemic problem of creating a learning economy is the design of the component institutions (e.g., corporations). This is especially important because so much learning occurs within organizations and because so much knowledge resides within firms. Chapter 2 described the marked differences in productivity across firms.

Somehow, knowledge that resides in the more productive firms does not seamlessly get transmitted and absorbed by the less productive.

Learning and the Boundaries of Firms

Within any institution, there may be incentives to develop knowledge and to hoard or transmit it. The issue of the architecture of a learning firm is parallel to that of the architecture of a learning economy. In some ways, the two cannot be separated. Traditional discussions of the boundary of firms (Coase 1937) focused on transaction costs, but equally important is the structure of learning. It may be easier to transmit information (knowledge) within a firm than across enterprises, partly because the "exchange" of knowledge is not well-mediated by prices and contracts.[1] If so, and if learning is at the heart of a successful economy, it would suggest that firms might be larger than they would be in a world in which learning is less important.[2]

While firms seek to maximize the flow of information/knowledge within themselves, realizing that knowledge is power (or at least money), they seek to limit the transmission of knowledge to others, for instance, requiring employees to sign nondisclosure agreements. Thus, firms go to great lengths to maintain secrecy. While for the advancement of society, it is desirable that knowledge, once created, be transmitted as broadly and efficiently as possible, profit-maximizing firms have traditionally sought to limit to the extent possible the transmission of knowledge.

The Design of a "Learning" and "Innovative" Firm

There is a large literature—too large to reference here—that describes how to "create" a learning firm. We describe some of the significant advantages that large firms have in financing research and bearing the associated risk. But large enterprises often develop bureaucratic structures to manage resource allocations, and while they prevent waste and even bad projects from being undertaken—reining in the excessive optimism that is often associated with the entrepreneurial spirit—those same processes may at the same time stifle innovation. So too, bureaucratic structures may reinforce the excessive loss aversion to which behavioral economists have called attention (see Kahneman 2011). And even if knowledge is *supposed* to flow freely within an organization, individuals realize that "knowledge is power" and attempt to hoard it, to impede its free flow. Innovative firms create organizational designs and incentive structures that attempt to address each of these problems.[3]

It is worth noting that a key aspect of this large literature on innovative firms is a notion that parallels one of the main themes of this book, on learning-by-doing—what is referred to as "learning by experience." (See, for instance, Morgan McCall's [2004] discussion of "Leadership development through experience" and the references cited there.)

Most of this book, though, is not concerned with how to design a firm to maximize learning, but rather how government policy can affect the structure of the economy to maximize societal learning. Accordingly, in the discussion that follows, we mostly abstract from microeconomic structures, focusing on broader policies: on the principles which should guide government intervention and on alternative instruments.

Why Industrial Firms Are the Source of Innovation

In the analysis in the ensuing chapters, a key assumption is that the industrial sector is the source of innovation. The justifications for such an assumption are rooted in the nature of industrial activity. Innovation activity takes place in firms that (relative to firms in other sectors) are (1) large, (2) long-lived, (3) stable, and (4) densely concentrated geographically. Agricultural and craft production, by contrast, typically takes place on a highly decentralized basis among many small, short-lived, unstable firms.

In the following paragraphs we describe in more detail why these attributes are conducive to learning, some of the reasons for the comparative advantage of the industrial sector in learning, and why that sector is more likely to give rise to learning externalities.[4]

(1) LARGE ENTERPRISES Since particular innovations are far more valuable to large organizations that can apply them to many units of output than to smaller ones with lower levels of output (see Arrow 1962b), there is far greater incentive to engage in R & D in large enterprises; because enterprises are larger in the industrial sector than in the agricultural/craft sector, there is likely to be more innovation in the industrial sector. The result will be higher investments in innovation in the former sector than the latter. This can be looked at another way: Large firms can internalize more of the externalities that are generated by learning.

Moreover, innovation is highly uncertain, and firms and individuals are risk averse. Large enterprises are likely to be less risk averse, and thus better able to bear the risks of innovation. Moreover, because of information imperfections, capital markets are imperfect, especially so for investments in R & D, which typically cannot be collateralized. Capital constraints are less likely to be binding on large enterprises.

(This does not necessarily mean, however, that within the industrial sector the largest firms are the most innovative. The difficulties of developing appropriate incentives for the reward of innovation may militate against large enterprises. There is an ongoing debate over whether large or small enterprises are most conducive to innovation. Large firms may have the resources to finance innovation, typically lacking in smaller enterprises, but there is an impressive record of large firms not recognizing the value of pathbreaking innovations, including Microsoft being too wedded to the keyboard, and Xerox not recognizing the importance of a user-friendly interface in its bid to bring its earlier computing innovations to consumers.)[5]

(2) STABILITY AND CONTINUITY The accumulation of knowledge on which productivity growth is based is necessarily cumulative. This, in turn, greatly depends on a stable organization for preserving and disseminating the knowledge involved and on continuity in jobs and personnel to support these processes. In large organizations, with the resources to provide redundant capacity where needed, the required degree of stability and continuity is much more likely to be present than in small, dispersed organizations, where the loss of single individuals may completely compromise the process of knowledge accumulation. As a result, steady productivity improvement will be much more likely to arise from industrial than agricultural and craft production.

There is another way of seeing why stability and continuity contribute to learning. As we noted earlier, the benefits of learning extend into the future. Long-lived firms can value these distant benefits more. Also, because industrial firms are typically longer lived and more stable than firms in other sectors, they can access capital at lower interest rates. They are likely to be less capital constrained, to act in a less risk-averse manner, and to discount future benefits less.

(3) HUMAN CAPITAL ACCUMULATION Opportunities and incentives for accumulating general human capital are likely to be far greater

in large, complex, long-lived, stable industrial enterprises with a wide range of interdependent activities than in small, dispersed, narrowly focused agricultural or craft enterprises. There is a greater likelihood of benefits from the cross-fertilization of ideas.

Long-lived stable firms have a greater incentive to promote increased human capital that leads to greater firm productivity and a greater ability to adapt to changing circumstances. As we noted in the previous paragraph, they also have a greater capacity to finance these investments and an enhanced ability and willingness to bear the risks. The resulting human capital accumulation is a critical element in both developing the innovations on which productivity growth depends and in disseminating them as workers move within and between enterprises and across sectors.

(4) CONCENTRATION AND DIFFUSION OF KNOWLEDGE ACROSS FIRMS Diffusion of knowledge among densely collocated, large-scale industrial enterprises (often producing differentiated products) is likely to be far more rapid than diffusion of knowledge among dispersed, small-scale agricultural or craft enterprises. The fact that they are producing different products enhances the likelihood that they will make different discoveries. The fact that they are producing similar products enhances the likelihood that a discovery relevant to one product will be relevant to another.

Recall that earlier we emphasized the importance of the diffusion of knowledge and stressed the key role that geographical proximity plays. More recent discussions of the role of clusters have reemphasized the importance of geographical proximity (see Porter 1990). Geographical proximity promotes cross-firm mobility, an important catalyst to learning and an important way in which learning gets transmitted. (Firms try to restrict the mobility of their workers out—it reduces the return to their investments in human capital and may reduce their competitive advantage over other firms, as the firm's knowledge is shared with others; at the same time, they try to encourage the mobility of other workers to their firm.

(5) CROSS-BORDER KNOWLEDGE FLOWS While learning is facilitated by geographical proximity, especially developing countries (where many firms are operating far below "best practices") can learn from advances in other countries. While agricultural conditions may

differ markedly from one country to another, the potential for cross-border learning may be greater in the industrial sector. The existence of large, stable enterprises with the incentives and capacities to engage in cross-border learning enhances the role of that sector in societal learning. Indeed, it is widely recognized that success in the industrial sector requires not just knowledge but also the ability to acquire knowledge that is common across borders.

Why There May Be Significant Spillovers from the Industrial Sector to the Rest of the Economy

Learning by one firm or subsector spills over to other firms and subsectors within the industrial sector, through, for instance, the movement of skilled people and advances in technology and capital goods that have cross-sector relevance. But the benefits spill over more broadly, even to the agricultural sector. In the following paragraphs we describe some of the ways that this occurs, especially as a result of the tax revenues that a growing industrial sector can generate. Large-scale, densely concentrated activities are by these very attributes far easier to tax than small-scale, dispersed activities.

(1) THE ABILITY TO SUPPORT PUBLIC RESEARCH AND DEVELOPMENT One of the important uses to which tax revenues can be put is to support publicly sponsored R & D. This factor may be especially important in the support of agricultural research, like that undertaken by the Agricultural Extension Service in the United States. These activities directly contribute to agricultural productivity growth but could not be supported without a taxable base of industrial activity. And the private sector, on its own, would not have undertaken this research, let alone the widespread dissemination which was so critical to the growth of agricultural productivity.

(2) PUBLIC SUPPORT FOR HUMAN CAPITAL ACCUMULATION Another important use of these funds is investment in human capital. Our earlier discussion focused on investments in human capital by the firm. These investments are, to a large extent, complementary to investments in human capital that occur within the formal education system. Just as in the case of R & D, private capital market failures mean that

public support in the form of free primary and secondary education is a critical component of general human capital accumulation and yields high social returns. Moreover, the high returns to education in the industrial sector lead to a greater demand for an educated labor force. And again, as workers migrate across sectors, ultimately higher productivity growth in the agricultural and craft sector will be engendered as well. Moreover, eventually public education gets extended to all parts of the economy, not just urban areas.

(3) THE DEVELOPMENT OF A ROBUST FINANCIAL SECTOR In chapter 3, we noted that the spillovers from one sector to another may take many forms. Much of the discussion focuses on technological spillovers. But there are important spillovers from the institutional developments that are necessary to make an industrial economy function. The previous paragraph stressed the benefits to the rest of the economy from the development of a general publicly supported education system. The heavy investment of a modern industrial economy requires finance. It is not surprising, then, that an industrial environment should be characterized by a more highly developed financial sector than that found in an agricultural/craft environment. Once developed, a strong financial sector facilitates capital deployment throughout the economy, even in the rural sector.

The above discussion highlights both why learning (innovation) may occur more rapidly within the industrial sector and also why the learning and innovation (broadly understood, to include institutional innovations) developed there have spillovers, not only within the industrial sector but to the rest of the economy. These spillovers involve knowledge, human capital, and institutional development. For instance, improvements in financial institutions and education have systemic benefits. While knowledge moves more freely within a country than across countries, it moves more freely across borders in the industrial sector, but some of the knowledge thus transmitted—including learning to learn—may be of value to other sectors.

There are still other channels through which spillovers occur. For instance, industrial mechanization improved agricultural productivity.

There are, in short, multiple channels through which the technological and institutional innovations of the industrial sector get translated into higher productivity growth for the economy as a whole.

Geographical Localization

The analysis in later chapters makes three critical assumptions: (i) the industrial sector is more successful in learning; (ii) spillovers are concentrated within national boundaries; and (iii) the learning spillovers across sectors within a country are significant—more significant than the cross-border spillovers. We have discussed at length the rationale for the first and third assumptions.

The second assumption rests on four factors: (1) geographic proximity, (2) international restrictions of the movement of labor (and associated movements in knowledge and human capital), (3) language and cultural barriers, and (4) historical patterns of social interactions, which are strongly affected by national boundaries and which get reflected not just in language and culture but also in transportation systems, social networks, and institutional arrangements. Individuals are the main carrier of learning, and the factors listed above imply that, by and large, labor mobility is easier within a country than between countries. Moreover, as we noted earlier, learning is local, and much of the knowledge which is relevant in one country may be less relevant in others.

The results of our analysis in later chapters, however, require only that transmission of knowledge to the agricultural and craft sector be stronger within a country than between countries. Indeed, our results are strengthened if there is some element of transmission across countries within the industrial sector, so long as that transmission increases with the size of the industrial sector in the developing country. For then, for the developing country, there is a further reason for promoting the industrial sector: It is the "window to the world," the channel through which more advanced knowledge gets transmitted to the developing country for both industry and agriculture. A manufacturer of textiles, for instance, absorbs information about textile production from other countries (perhaps because he buys machines from other countries). But some of that knowledge may be relevant for the agricultural or other sectors of the economy.

2. Macro-Conditions for Creating a Learning Society

Most of this book is concerned with microeconomic policies. But creating a learning society requires creating an economic environment that is conducive to learning. For this, the macroeconomic environment is central. Stability is important to the learning process.

Evidence for this comes from the experience of developed economies during recessions. Productivity growth is normally low during contractions, and there is no offsetting gain during subsequent expansions.[6] The productivity loss during the dislocation associated with the recession appears to be permanent.[7]

There are several reasons why stability is important for *societal* learning. The first is that much information is embodied within existing institutions, in complex webs of interactions. Key institutions—firms—often die in the face of high levels of instability. They are not simply brought back to life when the economy recovers. There are important hysteresis effects. New institutions may be created, but much of the knowledge that was embedded in the old institutions (especially knowledge that is typically referred to as "tacit") has to be re-created.

Contrary to the popular impression, it is not just the worst firms that die. For instance, during the East Asian crisis, the Korean firms that died differed little from those that survived—except that they were more encumbered with debt. Their main mistake was committed by the financial officers (who undertook excessive debt) and was predicated on (ex post) excessive confidence in macroeconomic stability. In short, though the evolutionary process showed itself to be enormously destructive (with 50 percent of all firms facing bankruptcy), it was not very creative.[8]

Moreover, managerial attention is limited. When firms are focusing on survival, they have less attention to devote to "learning," except learning how to survive.

Third, high levels of macro-instability lead firms to act in a more risk-averse manner. When firms go into recessions, among the first things to be cut are investments in R & D, and this is even true among firms that are relatively dependent on innovation. Part of the reason is that learning is future oriented. One has to make sacrifices today and undertake risks today for future benefits. But in the presence of instability, there is a risk that there will be no future—and hence less reason to make the requisite investments today. Instability weakens future-oriented incentives.

Fourth, learning requires resources, including access to capital. Instability may make capital less accessible and more costly.[9] In downturns, capital is likely to be rationed, and investments in R & D are often sacrificed (Greenwald, Salinger, and Stiglitz 1990).[10] Because investments in research cannot be collateralized and because they are particularly risky, they may be particularly hard hit in economic downturns.

Deep downturns have even more adverse effects, since in that case, not only are the balance sheets of firms hit but so are those of financial institutions. Hence, there can be system-wide constraints on credit availability, even for working capital, forcing firms to cut back even more on their investments, including and especially investments in innovation.[11] Similarly, cutbacks in production mean that there will be less learning-by-doing.

Thus, finally, instability affects the average level of output (that is, the average "output gap"—the difference between what the country could have produced and what it actually produces—is larger) and the structure of output. Capital-intensive industries, with larger learning potentials and more learning spillovers, may be disadvantaged relative to other sectors with less learning and less learning spillovers.

Thus, our view is markedly different from that which sees recessions as having a purging effect on the economy—the idea that the silver lining in the cloud of a recession is the shake out that results. In that view, firms cut out fat, fire unnecessary workers, and restructure the firm to make it leaner and meaner. Firms that are less efficient, that have been surviving off previously earned capital, can no longer do so. The Darwinian struggle for the survival of the fittest means that those firms that are less fit don't make it through a serious downturn, and thus recessions speed the process of natural selection. Schumpeter trumpeted these virtues. He wrote, on recessions:

> They are but temporary. They are the means to reconstruct each time the economic system on a more efficient plan. But they inflict losses while they last, drive firms into the bankruptcy court, throw people out of employment, before the ground is clear and the way paved for new achievement of the kind which has created modern civilization and made the greatness of this country.
>
> —SCHUMPETER 1934, 113

Schumpeter's views were much more in line with those of Andrew Mellon, secretary of treasury under President Hoover, who famously said:

> Liquidate labor, liquidate stocks, liquidate farmers, liquidate real estate . . . it will purge the rottenness out of the system. High costs of

> living and high living will come down. People will work harder, live a more moral life. Values will be adjusted, and enterprising people will pick up from less competent people.

Indeed, Schumpeter seems to have believed that, net, depressions were good for the economy, more in the nature of "a good cold douche" (Heilbroner 1980, 311).

While there may be some virtues in the process of "creative destruction" that is associated with innovation, the destruction that occurs in the process of cyclical fluctuations is not offset by any creation, and the anticipation of future volatility dampens investment in learning and R & D. The long-term benefits from the purging effect—the incentive to reduce slack posed by a recession—are outweighed by the costs associated with cutbacks in R & D. Part of the reason for this is that because of the learning spillovers upon which we focus here, in general, private firms underestimate the benefits of R & D; and because, especially in a recession, the social costs of unemployment are high, firms underestimate the social costs of "purging" and thus take it too far (Stiglitz 1994b). In short, volatility is bad for the long-term growth of the economy.

Policy Implications

This has important implications for policy: Policies which *expose* countries to a high level of instability or which increase the economy's instability (e.g., by weakening automatic stabilizers) have an adverse effect on learning. Examples of policies that have exposed countries to greater risk include financial and capital market liberalization and deregulation.

By the same token, policies that focus on price stability, at the expense of *real* stability, may be counterproductive (see Stiglitz et al. 2006). Inflation targeting, with its focus on price stability attained by interest rate adjustments, may be "doubly" bad: Responding to inflation by increasing interest rates—even when the cause of the inflation is an exogenous supply shock—is an example of a pro-cyclical policy. And the increases in interest rates have a disproportionate effect on certain sectors, those that are most interest sensitive and which rely most on bank financing. Small businesses, in particular, bear the burden. We have already noted that firms that die in a downturn don't come back to life when the economy recovers. There is a loss of informational capital.

So too, firms that may be killed when interest rates are raised dramatically don't come back to life when they are subsequently lowered. This is especially important in developing countries, where there may be a dearth of entrepreneurship. If, as some claim, much of the learning and innovation in society occurs within small and young enterprises, then these policies increase the burden on these key "learning" sectors. But whether that is the case or not, these policies exacerbate the already adverse effects on learning and investments in R & D arising from the cyclical volatility in the "shadow" cost of capital.

CHAPTER FIVE

Market Structure, Welfare, and Learning

ADVOCATES OF free market economies often stress the virtues of a market economy in promoting innovation. Remarkably, in spite of the constant praise of the market system's "innovativeness," there appear to be no general theorems on the efficiency of markets with respect to the pace and direction of innovation.

Joseph Schumpeter (1912, 1943), who argued for the centrality of innovation, cast aspersions on standard economic theory and its policy prescriptions, which had lauded the virtues of competition and castigated monopolies as the "supreme evil."[1] To Schumpeter, not only was the heart of capitalism innovation, but innovation required some degree of monopoly power.[2] If competition were perfect, and knowledge flowed freely, innovators would not be able to appropriate any returns to their innovations, and without innovations, economies would stagnate. Moreover, competitive firms would have difficulty raising the finance that investments in innovation require. Monopolies could generate the profits necessary to fund research—especially important in an era when financial markets were less developed and venture capital firms did not exist. Borrowing to finance speculative research was limited because, if the research project failed, there was nothing for the lender to seize.[3] In real estate, at least there is some collateral.

Schumpeter clearly took a more benign view of monopolies than did the rest of the economics profession. Indeed, he looked with some jaundice at the conventional economists' single-minded preoccupation with the dangers of monopolies and their hagiography of competitive markets. Schumpeter thus countenanced monopoly: He saw the monopolist's constraints on output as a small price in static inefficiency to pay for the speedier innovation that he thought was associated with monopoly.[4]

Writing after a period in which the capitalist economies' performance was less than stellar—the Great Depression, in which large fractions of the capital and human resources had been left idle, at great human suffering, for an extended period of time—Schumpeter could still look at the great sweep of history. Such fluctuations had happened repeatedly, and even taking the loss in output during such episodes into account, he noted the huge increases in living standards which capitalism had brought about and which it was likely to continue to bring about.[5] He was even optimistic about the elimination of poverty: With little evidence of an increase in inequality,[6] as average incomes were increased, it was likely that those at the bottom would see newfound prosperity.

Schumpeter did think competition was important, but the kind of competition that he thought was central was markedly different from that modeled in standard competitive theory, in the models growing out of the work of Walras, later to be refined by Arrow and Debreu. In those models, there were a large number of firms in each market, so many that each took the price it received for its goods as given. Schumpeter replaced the notion of competition in the market with competition *for the market*, argued for the benefits of the kind of creative destruction that arose out of the innovative process, and suggested that the monopoly power thus created would only be temporary. One monopolist succeeds another, and the threat of competition induces the incumbent monopolist to engage in a high level of innovation.

This and the next chapter address several related questions: (a) Does more *market* competition (that is, more competition *within a market at a particular point of time*) result in more innovation? (b) Was Schumpeter correct in his analysis of Schumpeterian competition, i.e., that competition *for* the market would sustain a high level of innovation? (c) Are markets efficient in the level and nature of the innovation and learning?

This chapter focuses on the link between competition and innovation, a link which is controversial and remains unsettled. The traditional view is that competition provides a spur to innovation, while monopolies are lethargic (see Leibenstein 1966). But Schumpeter, as we noted, saw the traditional economists' view as a fetish. Indeed, as we explained in the last chapter, large enterprises have distinct advantages in learning and innovation.

This chapter is divided into two parts. In the first, we analyze the market structure with innovation. We explain why, consistent with Schumpeter's insights, sectors where innovation is important are likely to be characterized by very imperfect competition. We also explain why Schumpeter's claim—that at any point of time there was likely to be a monopoly (or limited competition), but that such monopoly was only temporary—was wrong. Monopoly power is likely to persist. In the second part of this chapter we analyze the effect of market structure on innovation, explaining why the relationship is so ambiguous and why Schumpeter's claim that competition *for* the market was an important spur for innovation—more effective than competition in the market—is not, in general, correct.

This chapter primarily provides descriptive analytics. It simply compares (under certain idealized assumptions) the level of innovation under different institutional arrangements. It does not ask whether there is too much innovation under monopoly or too little under competition. That is the subject to which we turn in the next chapter.

1. Market Structure with Innovation

Markets with innovation are naturally not perfectly competitive. Investments in innovation are fixed costs. Consider a simple model where there are constant (marginal) costs of production. Those costs can be lowered by investing more in research; but the knowledge, once acquired, can be used no matter what the scale of production. If costs of production of one unit are lowered by a dollar, total costs are lowered by $1,000 if the firm produces 1,000 units, and by $1 million if the firm produces a million units. Because the value of any cost reduction is proportionately greater to larger firms, they have an incentive to "learn" more and engage in more R & D, giving them a competitive advantage over smaller firms. Over time, their cost advantages increase, to the point that they come to dominate the economy. We obtain a natural monopoly, so long as marginal costs are constant.

Even if one can write down a model in which there appears to be an equilibrium in which several firms operate (as Spence [1983] has done), the equilibrium is fragile and unstable. With learning by doing, even if the firms start out at exactly the same productivity, if one firm happens to sell or produce a little more than its rival(s), it will have a cost advantage, and the equilibrium with multiple firms quickly unravels, as described earlier.

Consider, for instance, a Nash-Cournot model and assume that the monopolist has a high discount rate—sufficiently high that it ignores the benefit of future learning. And assume that one of the firms has infinitesimally lower marginal costs than the other. It follows that it will produce more. But that means that it will learn more. And that means that in the following period, its level of production will be relatively greater, compared to that of its rival. Over time, one of the firms has an increasing competitive advantage over the other, until the firm that began with just an infinitesimal advantage becomes dominant—it becomes the monopolist.

Of course, if the firms are less myopic, then the firm with the competitive advantage realizes that it will be growing relative to the other, and that implies that it has an incentive to expand production now even more, relative to its rival. It would appear that convergence to the monopoly equilibrium would be even faster.

With Bertrand competition, divergence is still more rapid, for the firm with the slight competitive advantage grabs the entire market; it learns, and its rival does not.

Thus, the situation where identical firms might cohabit the market is a knife-edge equilibrium, unlikely to persist.

These results hold even if there are spillovers, so long as spillovers are imperfect. For as long as that is the case, the firm engaging in learning (R & D) will have a cost advantage over rivals that engage in less learning (R & D). In practice, there are likely to be some imperfect spillovers, so that markets where innovation is important are likely to be marked by large externalities and high levels of imperfections of competition. Thus, *both* the level of production and the level of innovation will be distorted.

The only way that there can be effective competition in the situation just described (constant marginal costs of production) is that there be full spillovers to others in the same industry; but if that is the case, then each firm will free ride off the research efforts of others and, if the number of firms is large, there will be no incentive to engage in R & D.

In the analysis that follows, we make two important distinctions. The first concerns the structure of the product market, the second the nature of the spillovers. In the discussion in the following chapters, we use a general formulation that has as limiting cases no and perfect spillovers.

As we have noted, endogenous learning makes some market structures infeasible: In the absence of full within-industry spillovers, a natural monopoly may exist. The market might then best be described as monopolistically competitive, with only one firm in each "industry" but with spillovers to other industries. In this case, we will see two distortions: underproduction as a result of the exercise of monopoly power and underproduction as a result of failing to take into account the learning benefits that accrue to others.

With full spillovers, there can be many firms in the industry; there will be competition, but as we have noted, there will be no investments in research.

Table 5.1 outlines the three cases. Much of our attention will focus on the cases with perfect competition and full spillovers, or monopolistic competition and no spillovers. As we have explained, the only case which is consistent with competitive markets (in the absence of some other source of strong decreasing returns to scale) is full spillovers.

The arguments that we have given for why markets with learning are likely to be dominated by a single firm also apply for markets where innovation arises from R & D. Here, the question of exclusivity becomes central; does the innovator have exclusive rights to the technology, or can others also use the technology (just as others "learn"

TABLE 5.1
Spillovers and Market Structures

	No Cross-Firm Spillovers	Full Cross-Firm Spillovers
Perfect Competition	X (not feasible)	Underinvestment in learning
Monopolistic Competition	Restricted output	Both market distortions

from the advances of the leading firm)? While in many cases, it is patents that give rise to exclusivity, there are other ways that the first firm to innovate can come to dominate, if not the market, at least the new technology. There is a first mover advantage; the first firm can succeed in developing loyal customer relations, establishing a reputation as the innovator. His higher level of production gives him a learning advantage, a head start, which others will be hard pressed to overcome.

Why There May be Some Competition Even in Highly Innovative Markets

Schumpeter seemed to agree with the view that the "natural" market structure at each moment was a monopoly. While there are some instances in which this is the case, there are many highly innovative markets in which it is not. Understanding why this is so is important if we are to understand the effects of competition on innovation.

Competition can, of course, be sustained if there is some offsetting force for decreasing returns (e.g., diseconomies of scale arising from, say, limits to managerial span of control). But in sectors in which learning is important, even then there are likely to be only a limited number of firms, and thus very imperfect competition. The extent of competition will depend on the importance of decreasing returns, on the one hand, and the extent of spillovers, on the other. When there are large spillovers (from the technological leader to others), then the other firms can still provide an effective challenge to the leading firm.

Even with patent exclusivity in controlling a technology, patents do not lead to a single firm dominating the product market. Those with the prior technology may still be able to compete, though with a cost (or product) disadvantage. This still can greatly circumscribe the innovators' monopoly power. (More generally, there are typically products which are partial substitutes for the product in question.)

Moreover, there may be alternative technologies for producing the same product. Even if Xerox had a sufficient hold on patents associated with the Xerox technology, there are alternative technologies for photocopying. Improvements in these alternative technologies eventually ensured effective competition for Xerox.

In addition, it is often possible to innovate around a patent. Pharmaceuticals are an industry in which patents have played an important role, yet particularly with very valuable drugs, companies have found "me too drugs" that have similar therapeutic effects but do not conflict with the patent. In some industries, like some types of metallurgy, it is so sufficiently easy to invent around a patent that patents do not play a major role.

Thus, Schumpeter's characterization of markets being characterized by a succession of monopolists is overly simple. There may be some products for which this is true, but, more generally, there is both competition in innovation and competition in the product market. And for reasons that we have already explained, the two are closely linked: firms that are producing are in an advantageous position for innovating. Still, analytically, it is useful to separately analyze the effects of competition in the market from the effects of competition for the market (innovation competition), and we do this in the discussion in the second half of this chapter.

The Temporary Nature of Monopoly Power

Schumpeter was not much worried about the domination of the market by a single firm at any point of time. He believed, as we noted, that that domination would only be temporary and that the fight to be the next monopolist would spur innovation.

While Schumpeter was correct, as we have seen, that at each moment of time markets may be dominated by a single firm, monopolies can be far less temporary than Schumpeter thought; while in some cases the threat of entry can be an important impetus to innovation, in other cases, to maintain their monopoly power, firms devote considerable resources to creating socially unproductive entry barriers. In doing so, incumbents can discourage the overall pace of innovation. Microsoft has become the poster child of how an incumbent can discourage innovation. Modern monopolists have become highly innovative—in creating new forms of entry barriers and in extracting rents out of their monopoly power.

By controlling the "platform" (PC operating system), Microsoft could use anticompetitive practices to leverage and extend its market power in the operating system to other areas, and in so doing undermine, and in some cases drive out, potential innovative competitors: Netscape in the browser market, Real Networks in the multimedia market. By bundling, for instance, Internet Explorer into the operating system, it was charging in effect a zero price—and it is hard to compete against a zero price. It was clear that a zero price did not maximize Microsoft's short-run profits, but the company seemingly believed that (if it could get away with it) such predatory behavior was consistent with long-run profit maximization. (Indeed, it turned out that even after it was found guilty of anticompetitive practices and some of these practices were enjoined, Netscape did not revive. Effective competition was brought about only by Mozilla's open-source Firefox.[7])

Dasgupta and Stiglitz (1980b) established a general result showing that monopolists have not only the ability but the incentive to preserve their monopoly power by doing more research than competitors and thus preempting them.[8]

The reason the monopolist will always preempt potential competitors is intuitively clear. If any other firm were to win the patent, the industry would be characterized by a duopoly structure. However, the existing monopolist can always ensure that it remains a monopolist by spending a little more on R & D than any potential competitor would find profitable. It is always in the monopolist's interest to do so, because by remaining a monopolist it can earn a flow of profit in excess of the sum of the two firms' profits under duopoly.

The argument is, of course, reinforced if the existing monopolist is more efficient than its competitors in R & D activity or in ancillary services (e.g., in distribution). One of the implications of the analysis of this book is that incumbent firms do have a knowledge advantage. Thus, if there is competition in R & D activity, there are strong tendencies for a monopolized industry to remain a monopoly. The fact that the monopolist is threatened by potential competitors at most spurs the monopolist to spend more on R & D than it would otherwise. But (at least in this model), the industry remains a monopoly.

The results are in many ways far more general than this model suggests. There can, for instance, be uncertainty in the research process, both about time and the eventual cost of production; but if there is any research project which is worthwhile for an entrant to undertake, it is profitable for the incumbent to preempt.

Why Contestability Doesn't Suffice to Ensure Innovation, Efficiency, or Zero Profits

Even if the monopolist perpetuates itself, it is possible that the threat of competition spurs innovation—forcing the level of innovation to be so high that profits are actually zero (so that, in the absence of government subsidies, there could not in fact be any higher pace of innovation)—or so the advocates of Schumpeterian competition claim. The notion that all that is needed for markets to be efficient is that there be competition for the market (potential competition), rather than competition in the market, is obviously very appealing and applies to all natural monopolies, not just those that arise out of the fixed costs associated with innovation. Were it true, it would mean that government wouldn't even have to worry about regulating natural monopolies. Potential competition would provide the necessary market discipline. Potential competition would ensure that the monopolist not exploit its market power by charging a price in excess of average cost. In the context of an innovation economy, it would mean that the monopolist would be forced by potential competition both to innovate and not to exploit its (temporary) monopoly power by earning excess profits.

This idea gained currency at the time that the AT&T monopoly began to be attacked in the 1970s. The defenders of monopoly argued that potential competition would ensure that profits would be driven down to zero. Markets where potential competition was sufficiently strong to do this were called *contestable*.[9]

It turns out, however, that the notion of contestability is not robust. Even if there are arbitrarily small sunk costs, even very strong potential competition will typically not result in zero profits. This can be seen in a simple example, where there are small sunk costs and fixed marginal costs. Assume the entrant is just as efficient as the incumbent and that after entry, the two engage in Bertrand competition (i.e., each takes the price of the other as given). Then after entry, the price will be driven down to marginal costs; there will be no profits. But that means that the entrant will lose its sunk costs. Knowing this, the entrant will not enter. The incumbent can charge the monopoly price with impunity.

What matters is not, of course, the level of competition now, but that after entry. And if there is keen competition after entry, then potential entrants will be deterred from entering. Ironically, the stronger competition is in markets after entry, the less likely that there will be entry and the more likely that a monopoly will be sustained (Stiglitz 1987c; Farrell 1986).

Since R & D expenditures (or investments in learning) are by their very nature sunk costs, and in the sectors with which we are concerned, where R & D or learning is important, they are significant, it should be clear that the contestability notion is of little relevance; potential competition does not suffice either to ensure efficiency or zero profits (see, in particular, Dasgupta and Stiglitz 1988b).

Entry Deterrence

The analysis of the previous section argued that potential competition does not ensure effective competition or efficiency—prices may remain sustained well above marginal costs, and potential competition will not force the incumbent to engage in the efficient level of research. But matters may be even worse, for firms may undertake costly actions to deter entry (Stiglitz 1981, 1987c). They may, for instance, hold excess capacity, so that potential entrants will know that there can be a strong competitive response to entry. Or they may invest in high levels of cost reduction to preempt rivals from entering—levels that are greater than the most efficient or socially optimum.

Not surprisingly, efforts at entry deterrence can be welfare decreasing (Stiglitz 1981). In particular, they can discourage innovation. A potential innovator knows that a dominant firm like Microsoft can engage in predatory and other behavior that will result in the innovative firm not capturing much of the benefits of its innovation.

Why Patent Races May Provide Only a Limited Spur to Innovation

In the context of innovation—patent races—entry deterrence may be even easier than the above static analysis suggests. Patent races are dynamic. The first firm to patent gets a big prize; the losers get nothing—or at least very little. (The knowledge that they acquired may have considerable value, enabling them to compete more effectively in some other race. For purposes of simplicity, we will ignore this in the subsequent discussion.)

One of the problems of contests in general is that if there is some firm (individual) that is viewed as likely to win, others will drop out of the competition. Why put up the effort if the chances of winning are low? But if this happens, contests may be much less effective in spurring effort than is generally assumed. Nalebuff and Stiglitz (1983b) studied equilibrium behavior in contests (with possibly many contestants). They show that contests have to be carefully designed to avoid this pitfall. For instance, one way of ensuring that all participants exert effort is to punish the loser (the one in the last place), rather than rewarding the winner. But in the context of innovation races, we can only tell who is the winner—the person who makes the discovery.

Thus, in a patent race, the optimal strategy of the incumbent is simply to get far enough ahead of the rivals that they are discouraged from entering—and having done that, it can rest on its laurels. Patent races may provide only limited spurs to innovation (Fudenberg, Gilbert, Tirole, and Stiglitz 1983).

(Matters may not be as bad as this analysis suggests, since there are often multiple dimensions to [product] innovations. One innovator may succeed in getting a product that is better in one dimension, another in another. What matters is that one is not dominated in all dimensions. Those who fail in every dimension are the "losers," and market competition may drive them out of business. Thus, the force of competition may serve to encourage high levels of research by punishing the laggards, consistent with the Nalebuff-Stiglitz analysis.)

Some Caveats on the Ability of a Dominant Firm to Perpetuate Dominance

There are three caveats to these conclusions suggesting that a firm, once it becomes dominant, will remain so. The first concerns diseconomies of management: the monopolist, which has to manage the task of production as well as innovation, may become less nimble in its research. If the entrant has, as a result, a cost advantage, then it may not pay the incumbent to preempt its rival. This is especially so if there are multiple avenues to be pursued. If the firm attempts too many diverse research projects, performance will deteriorate.

Second, the rival may be irrationally enthusiastic about its abilities, and this may lead it to engage in an excessive level of research.[10] Given this irrational exuberance, the level of R & D that the incumbent would have to undertake to preempt the rival may be very high—so high that the incumbent may decide not to preempt. This is especially so because there is a great deal of uncertainty about the outcome of R & D.

Alternatively, the entrant may engage in a sufficiently low level of R & D that the incumbent thinks that there is a low probability of the entrant being successful enough to become a major player. (It may in fact share some of the same overconfidence in its relative abilities.) Given this low probability, it simply doesn't pay to engage in even more

R & D to drive the entrant out. But even at the low level of research, there is a *chanc*e that the entrant will be successful—and the incumbent will not be—so that the entrant displaces the incumbent as the dominant player.

Finally, there may be multiple technologies for producing the same goods or services. While the incumbent may have a comparative advantage in the current technology, there may be an alternative technology, related to a technology employed elsewhere in the economy; firms employing that technology may have a comparative advantage in lowering costs associated with that technology.

In other words, it may not pay the incumbent to explore every possible technology and preempt potential entrants in each of these areas; so long as that is the case, stochastically, one of the other researchers will, eventually, succeed in developing an alternative technology.

In any of these cases, there can be entry, and the new entrants can even displace the existing firm. But the lesson of this chapter is still relevant: monopolies can persist for a long time. While they persist, markets are distorted, and, as we shall see more clearly, the threat of competition does not, in general, induce the optimal amount of research and learning.

2. The Effect of Increased Competition on Innovation

The thrust of the previous section was that the degree of competition in the market was endogenous, and that there were many circumstances in which there might be both a limited number of firms (in some cases only one) producing and a limited number of firms competing to be the next monopolist. What can it mean, then, to ask what is the effect of competition (say the number of firms in a market and how they interact) on innovation? There are two thought experiments: We could ask, what is the effect of a policy which prohibited mergers that firms believe would enhance their profitability? The number of firms, in this case, would be greater than would occur in an unrestrained equilibrium, and it is meaningful—and important—to ask what the effect of such a policy on innovation is. Alternatively, we can ask what is the effect of, say, a change in the cost of innovation (for instance, as a result of subsidy for research). Such a change will have two effects: a change in the level of investment in innovation *at any level of competition (number of firms)*, and a change in the number of firms (and the level of investment of each resulting firm). The change in the pace of innovation as a result of this change in the number of firms is what we mean by the effect of competition on innovation.

Whether more competition leads to more or less innovation turns out to be a complicated question which depends on whether we are referring to competition in the product market (ex post competition) or competition among innovators (ex ante competition); the *nature* of competition in the product market; the nature of the innovative process itself (for instance, the riskiness of the innovative process); and whether there is *exclusivity* and the source of that exclusivity—whether the patent system gives the first to innovate at least temporary monopoly rights, or whether the first to innovate obtains *de facto* exclusivity as a result of its first mover advantage.

It is important to recall, as we emphasized earlier, that more investment in innovation does not necessarily translate into a higher pace of increases in standards of living—what we mean by societal innovation. The market innovation process can be highly inefficient. Patents can be used to block the real innovation of others—what are called hold-ups—and to extract rents from the "real" innovators. Innovations can be used to attempt to enhance market power—monopoly distortions—for instance, by extending the patent life (in a process called evergreening). Innovation can be used to circumvent regulations designed to ensure the stability and efficiency of the economy (arguably, much of the innovation in the financial sector was of this form).

The previous chapters have shown why competition in the product market and competition in the "innovation market" are interconnected: one learns from producing, and, in the process of production and marketing, one learns much about what are likely to be the most profitable avenues of exploration even if knowledge is not just the byproduct of production but the result of the explicit allocation of resources to R & D. This is one of the reasons that market dominance persists.

While the two forms of competition are interlinked, analytically, it is convenient to separate the effects of more competition *ex ante* from the effects of more competition *in the market*.

Effects of Ex Ante Competition[11]

Increased ex ante competition has several effects which can lead to a greater or slower pace of innovation. Competition can affect both the ability and willingness of firms to engage in research, and the outcomes of the research process.

A DIVERSIFICATION EFFECT When there are many researchers pursuing different research strategies, it is more likely that at least one will succeed. Of course, the benefits of diversification are reduced if the different firms engaged in research follow very similar research strategies. And it is possible that a single firm pursue multiple research strategies. Still, the culture and mind-set of a firm is likely to induce it to think that a particular strategy (or a limited set of strategies) is most likely to succeed, and accordingly, a greater diversity of firms is likely to lead to a greater diversity of strategies.

A PRO-COMPETITION (CONTEST) EFFECT This is the effect seemingly most stressed by Schumpeterians: the competition to be the monopolist (if only a temporary monopolist) spurs each competitor to work harder, to beat the others. If in a patent race, there are significant marginal returns in terms of increased speed (a faster arrival time for the innovation) to investing more, then as each competitor tries to outcompete others, the pace of innovation quickens.

Note that with contests, the marginal return to investing more may be increased with more competition, even if the average (expected) return is decreased. This is one of the reasons that contests are such a powerful tool for enhancing incentives. We shall go on to note other instances in which marginal and average returns change in the opposite directions.

The magnitude or even the existence of this pro-competition effect has been questioned. As we noted, the incumbent firm may, for instance, be able to take actions which pre-empt *any* rival, in which case an increase in the number of (potential) rivals makes little difference.

AN AGENCY EFFECT The standard theory of monopoly holds that monopolists are efficient in their production decisions; the only distortion arises from restrictions on the level of production. In this context, we have assumed that if there is a monopoly, the monopolist chooses the level of innovation to maximize its profits. But there is a wealth of evidence and some theory that without the discipline of competition, monopolies in fact often become inefficient and lethargic. The managers of monopolies may rest on their laurels, enjoying their share of monopoly rents, rather than maximizing the present discounted value of monopoly profits. (See Leibenstein [1966], who refers to the loss of efficiency from what is called today agency problems as X-inefficiency and suggests that it may be far more important than the allocative efficiency that economists have traditionally focused upon. Hart [1983] provides a more formal analysis of the issue in the context of innovation.[12])

In the presence of imperfect information, it is hard to design good incentive structures to motivate managers. If things turn out badly, is it because the manager didn't work hard enough, or because there was an adverse turn of circumstances? If the firm fails to make the innovation that it sought to achieve, was it because it turns out that the problem was harder than was at first realized, or because the researchers didn't exert sufficient effort? In these circumstances, *some* competition may spur innovation and help in the design of good compensation schemes.

This effect can be especially important when moving from a single firm to two competing firms. Competition provides information on the basis of which firms can design better incentive contracts for their managers, inducing higher levels of innovation.[13]

AN ADVERSE INCENTIVE EFFECT While in some circumstances increased competition can have a positive effect on innovation, in others there can be an adverse incentive from additional competitors. With exclusivity, each firm knows that its chances of winning the patent race are reduced, and this normally reduces the marginal returns to investments in innovation, potentially significantly.

And in the case of non-exclusivity, if there is more ex ante competition, it is more likely that there will be more competition in the subsequent product market, and hence, even if the firm is successful in its innovation, returns will be lower.[14]

AN EX ANTE RENT-STEALING EFFECT Even when there is a single firm that is a temporary monopolist, competition limits profits, for if there is more intense competition in the innovation market, the length of time that the dominant firm remains on top will be shortened. If competition were only over the size of the next step (the quality of the successor product), and the time of arrival of the next product were fixed, then more competition for the market would not affect the profitability of the successful innovator. But more realistically, an increase in innovation competition is likely to shorten the period over which the innovator dominates, with the consequence that the returns to investments in R & D are lowered.[15]

THE COMMON POOL EFFECT: ADVERSE IMPACTS ON OPPORTUNITY SETS As important as, or even more important than, incentives in determining the level of investment is the opportunity set confronting potential researchers (see Dosi and Stiglitz [2014] and the literature cited there). Each innovator draws upon the common pool of knowledge and adds to that pool. When what is added to the pool is greater than what is taken out, the pool grows and innovation increases. Government basic research adds to the pool. Typically, a patented innovation reduces the set of ideas available for others to draw upon, though associated with many innovations are unpatentable ideas that can catalyze research and learning by others. An increase in the number of competitors in the presence of "strong" intellectual property rights (allowing for greater withdrawal of knowledge from the pool of available ideas) can, under plausible conditions, reduce the pace of innovation, even though *at any size of the pool* innovation is incentivized.[16] The effect from the diminution of the opportunity set dominates (Stiglitz 2014a).

FINANCIAL CONSTRAINTS Many firms do not have the resources to finance their research and may have difficulties getting access to finance from others. Much lending is based on lenders' judgments about the likelihood of success. With many competitors, the likelihood of success of any one competitor is reduced, and hence the access to finance of each may be reduced.

ADVERSE EFFECTS ON INNOVATION ARISING FROM INEFFICIENCIES ARISING FROM PATENT EXCLUSIVITY There are many dimensions to a firm's research strategy—speed (how fast on average it expects to get the innovation); size (how "big" an innovation it is striving for, e.g., how big of a reduction in the costs of production); and risk (it can follow a strategy to pursue a small step with a high probability of success, or a big step with a small probability of success). The set of feasible research strategies can be described as a probability of making an innovation with a cost reduction of Δc at time T, given the investment level I. The firm chooses, from among the feasible research strategies, that which maximizes its (expected utility of) profits, given its beliefs about the research strategies of its rivals. It should be clear that in this framework, we cannot even summarize the research strategy by a few parameters, say, the level of investment, the speed of research, and the size of the innovation; there is an entire probability distribution.

There are a large number of inefficiencies that are generated in an intellectual property regime, particularly so if it is poorly designed, as we note in chapter 12, and these may increase with the number of innovators, so that even if the total level of investment in innovation were to increase, the pace of societal innovation may decrease.[17] Most obviously, some (much) of the research may be duplicative, which is especially important if different firms pursue highly correlated research strategies. Research may be directed at stealing profits away from rivals (me-too pharmaceutical innovations), with little benefit to real innovation. In this case, more firms may lead to more investment, but again little more societal innovation. With many firms doing research on a complex set of ideas, all of which are related to the production of a complex product, the difficulties in bargaining to achieve an acceptable distribution of rents among all the owners of the different intellectual property components becomes greater. The patent thicket becomes more dense, and this discourages innovation. With even two claimants to relevant pieces of intellectual property, the development of a product can be greatly retarded, as evidenced by the arrested early development of the airplane in the United States.[18] More generally, the most important input into any research is the result of other researchers' efforts, and with many firms controlling this prior knowledge, the returns to follow on research may be lower (and riskier, since the returns may be the result of an uncertain bargaining problem).[19]

The race to be first may itself give rise to inefficiencies. An optimal research trajectory may not be the fastest trajectory. On average, one might save on (expected) costs by a more orderly research process. Such a process might maximize the present discounted value of expected benefits minus costs, but if there are substantial rents associated with such a research strategy, it will pay another firm to outpace this firm. More generally, there may be a sequence of excessively rapid, small innovations, when the overall pace of innovation would be maximized with fewer large innovations.

These various effects interact in complex ways. For instance, in one simple model where firms, by investing more, can enhance the earlier arrival of a project, the adverse effect of competition on incentives can be larger or smaller than the beneficial diversification effect. In one particular parameterization, it can be shown that there can be an inverted U-shaped relationship (moving from one to two firms significantly attenuates incentives, so much so that the pace of innovation is reduced; but when the number of competitors is increased beyond two, the diversification effect dominates the incentive effect) or a monotonic relationship, where the diversification effect always dominates the adverse incentive effect.

The discussion of the common pool effect highlighted that innovation may diminish with a "tighter" intellectual property (IP) regime, even if incentives to innovate *at any given set of opportunities* are increased, and the discussion of the adverse effects of the patent system highlighted that even an increase in investment in research by each firm may not lead to faster societal innovation.

Competition in the Market

Whether more competition *in the market* leads to more or less innovation also turns out to be a complicated question—the answer to which depends on some of the same factors identified earlier: the *nature* of competition in the product market; the nature of the innovative process itself; and whether there is *exclusivity* (e.g., the patent system gives the first to innovate at least temporary monopoly rights).

Here, there are four effects which are central.

SIZE EFFECT When there are many firms in a market, the market gets fragmented, and each firm therefore produces less. This means the benefit that it gets from reducing costs is lowered—and thus its incentives to innovate are attenuated.

This effect can be seen clearly in the standard Cournot model (in which there are a fixed number of firms, and each takes the level of production of its rivals as given, and therefore maximizes its revenues). An increase in the number of firms increases overall production, but the level of production of each firm diminishes, and with that, the level of innovation (with or without spillovers).[20]

But while there may be some presumption that that is the case, it is not inevitable. In the case of Bertrand competition, where each firm takes the price of its rivals as given, moving from a monopoly to just a duopoly can lead to more innovation. The reason is simple. Assume that the innovator has exclusive rights to his innovation, but that the innovation is a relatively small innovation, so that the prior innovation can still be marketed, say at its marginal cost of production. This limits the price that the innovator can charge. But because the innovator knows that his price will be limited by the existence of competition, he knows that the scale of output will be larger. Unlike with Cournot competition, with Bertrand, the existence of competition leads to *larger* production for the innovating firm, and thus to greater incentives for investments in innovation, and a higher level of innovation.

FINANCE EFFECTS What is striking about the Bertrand model just described is that while the marginal return to investments in innovation is increased, the average return is decreased—simply because the existence of competition has significantly circumscribed the ability of the innovator to exercise the monopoly power arising out of his cost (product) advantage.

The average level of profitability can be important in the presence of financial constraints. A lender, in assessing the risk of providing funds, is interested in whether the average expected returns will suffice to pay back the loan; and the decrease in average returns will therefore reduce access to credit and thus, potentially, the level of innovation.

RISK EFFECTS With exclusivity, an increase in competition increases the riskiness of investments in innovation (and with risk-averse firms, the level of investments in innovation and therefore the pace of innovation). There is a smaller probability that any given firm will win the "winner take all" lottery—and a higher probability that he will emerge with nothing.

Greater risk may also adversely affect the ability of cash constrained firms to obtain finance, and thus indirectly adversely impact innovation.

On the other hand, in the more general case of non-exclusivity, the nature of competition in the product market has significant implications for the riskiness of investments in innovation. With Cournot competition, a firm that succeeds in getting costs that are slightly lower than its rivals or an innovation slightly faster than its rivals will have profits that are slightly larger than its rivals, and slightly larger than they would have been if he were slightly less successful. By contrast, in Bertrand competition—in a winner takes all market—the firm that comes in second gets nothing, and how much the first firm gets depends not only on his success but on his success *relative* to his rival. If his costs are just a little lower than that of his rival, then his profits will be small.

But the way in which competition affects risk in each market structure also depends on the nature of the innovation process. Assume, for instance, that the difficulty of reducing costs depends on a common factor facing all firms as well as idiosyncratic factors facing each individual firm. Bertrand competition essentially eliminates the risk associated with the common factor.

EX POST RENT-STEALING EFFECTS VS. ADVERSE EFFECTS ON INSTALLED ASSETS Some of the profits that accrue to any innovator are rents that would otherwise have accrued to others. This is most obviously the case in the me-too innovations in the pharmaceutical industry referred to earlier, but it is also the case in any market with imperfect competition. One firm's gains are at the expense of others—the increase in profitability does not entirely reflect a gain in societal welfare.[21] This implies that the incentives to innovate for, say, a duopoly are (on this account) greater than for a monopoly. Indeed, a monopolist who has an installed base of assets will take into account the adverse effects of any innovation he introduces on assets that he owns. To put it another way—a monopolist that introduces a new product might make considerable profits on the product, but he will realize that at least some of those profits are at the expense of his other products. He is, in effect, "stealing" profits from himself. This obviously dampens his incentives for innovation. The result is that there may be a positive effect from at least some competition.

It would be nice if we could derive simple conditions under which one or the other effect that we have delineated dominates, but to do so does not appear to be possible. Indeed, all the effects (those associated with ex ante and with ex post completion) interact in complex ways so that the net effect depends on the specific nature of the industry (e.g., on the form that competition takes in the industry, the importance of knowledge spillovers, and the nature of the innovation process).

One general result highlights the ambiguities; there is even a central case where market structure has no effect on innovation. With Bertand competition, under quite general conditions the set of research projects undertaken is invariant to the number of firms (Sah and Stiglitz 1987a). Consider a case where there is a probability of any project being successful, and that if it is successful, costs of production will be lowered to the same level. The pay-off from a given project depends on whether it is successful and whether others are successful. If no other project is successful, the pay-off is just the value of the cost reduction; if some other project is successful, the *marginal* contribution of this project is zero. And this is true regardless of the number of firms undertaking the projects; it is even true if there is only one firm.[22]

A Brief Note on Endogenous Market Structure and the Relationship Between Competition and the Number of Firms

The earlier discussion made clear why, in general, it made little sense to ask what the relationship is between the number of firms in a market and the level of innovation: both are endogenous variables, as Dasgupta and Stiglitz (1980a) emphasized. An increase in, for instance, the opportunity set (holding everything else constant) would lead to both more firms and more innovation. At the same time, an increase in the fixed costs of entering the market and conducting research (holding everything else constant) would lead to fewer firms and less innovation.

Thus, the sometimes suggested inverted U-shape relationship (Scherer 1967), which says some competition is good for innovation, but too much is adverse, has little basis in theory, even though it has played an important role in the macroeconomic literature on Schumpeterian competition.[23] Goettler and Gordon (2014) highlight the lack of generality of the inverted U-shaped relationship, showing that the nature of the relationship depends on the explanation for the increase in competition. They obtain an inverted-U relationship when the degree of product substitutability is varied, but a U-shaped relationship when entry costs are varied. Similarly, we noted that there may be a U-shaped relationship in a simple model of Cournot competition with innovation, as we trade off the adverse effects of incentives with the positive effects of diversification.

Empirical studies attempt to control for "everything else," but seldom do so perfectly or even well, since critical variables, like the fixed costs of entry and, even more so, the nature of the opportunity set, are not easily observable, and the problems are not convincingly addressed by standard economic techniques. It is arguable, for instance, whether any of the empirical studies provide a structural model or adequately control for some of the structural properties which have been shown in theoretical models such as that of Goettler and Gordon (2014) to affect the relationship between competition and innovation. (See also Gilbert [2006].)

Since the level of competition is itself an endogenous variable, the question of relevance is not so much the effect of competition on innovation but the effect of particular policies—like the stronger enforcement of anti-trust laws or stronger intellectual property rights—on both innovation and consumer prices, on societal welfare both in the short run and the long. This is the question to which much of the rest of the book is devoted, and which we frame in broad terms in the next chapter.

Concluding Comments

One purpose of this chapter has been to show that the relationship between competition and innovation in quite standard models of market interaction is far different than has been widely presumed. Indeed, in a standard Nash-Cournot model, there is a presumption that more competition leads to less innovation.

There are other factors that affect the relationship between competition and innovation beyond those on which this chapter has focused. For instance, larger firms may have a greater ability to bear the risks associated with innovation and get access to (or generate) the funds necessary to finance it. We explained in the last chapter other advantages that large firms have in conducting research. At the same time, larger firms can become more bureaucratic and less nimble, and in sectors where the pace of innovation is very rapid, that may put them at a disadvantage.

There are other aspects of the structure of the economy and society which may affect the relationship between competition and innovation. Well-functioning capital markets, with robust venture capital firms, may be able to provide greater access to finance for small firms. A stronger welfare state, absorbing some of the risks that individuals face, may enable even smaller firms to undertake riskier projects.

This chapter has provided some support to Schumpeter's view that markets dominated by a single firm may be more innovative than more competitive markets.

But Schumpeter was too sanguine about Schumpeterian competition. He overestimated the force of competition "for the market" as a spur to innovation and underestimated the ability and incentive of the incumbent monopolist to deter entry and to maintain its monopoly position.[24] These are not just theoretical niceties describing what might possibly happen; Microsoft took actions to discourage and suppress potential rivals so as to maintain its dominant position. Its predatory behavior lowered the returns that these innovative rivals obtained on their inventive activities, serving notice to other potential rivals that Microsoft was able and willing to engage in activities to discourage entry—even if they were flagrant violations of competition laws, and even if they entailed significant short-term profit losses. In doing so, the pace of innovation was almost surely lowered both from what it would have been in the absence of this anticompetitive behavior and from the socially optimal level.

We have seen that the relationship between the level of competition (both *for* the market and *in* the market) and innovation is complex. When, for instance, the research process is stochastic with success of different firms being imperfectly correlated, then the more firms (the more research projects) the greater the likelihood of success *if the effort of each enterprise were fixed*. But whether competition *spurs* innovation or dampens it (given that the chance of becoming the dominant firm is diminished) is ambiguous. While we explained how competition can enable the design of better incentive structures, more competitive patent races (races with more entrants) may not lead to more innovation. A well-designed contest with just two contestants may, in some circumstances, be optimal. There may be an inverted U-shaped relationship. But, by contrast, in other circumstances, there may be a U-shaped relationship between competition and innovation.

The result derived here that, in standard models, the level of innovation *may* be higher with monopoly than with duopoly, or than with more competitive market structures more generally, runs at odds with a strong presumption that competition is good for innovation. But the failure of monopoly lies, we suspect, outside the bounds of this model (and other standard models). Perhaps most importantly, the standard theory of monopoly holds that monopolists are efficient in their production decisions; the only distortion arises from restrictions on the level of production. The monopolist chooses the level of innovation to maximize its profits. Monopolies, unconstrained by competitive threats, still have an incentive to innovate. The (expected) present discounted value of profits is increased as a result of cost-reducing innovations, or as a result of the creation of products that consumers value more. Our analysis noted the limited impact of the pressure of competition *in standard models*.

But we also explained how lack of competition may lead to monopolies becoming inefficient and lethargic. Thus, the real reason that more competition may lead to more innovation is not captured well in standard economic models. The failure of monopoly is related to problems associated with agency costs and the design of good incentive structures.

This chapter has highlighted the ambiguous relationship between competition in the market and competition for the market and the pace of innovation. But even if more competition, either ex ante or ex post, were to lead to more *investments* in research, it does not mean that the market would maximize societal welfare, or even the pace of social innovation. As we have noted, the investments in research and learning may not be directed at improving societal well-being, and even when they are, the decentralized market-driven innovation process may be far from efficient.

Before turning to a detailed analysis of how the government might improve societal welfare by helping create a truly creative learning economy later in this book, we investigate in the next chapter in more detail the various reasons that the market economy misallocates its scarce innovative resources and the ways in which it does so. In doing this, we further explicate why the government has a central role in creating a learning society.

CHAPTER SIX

The Welfare Economics of Schumpeterian Competition

SCHUMPETER AND his latter-day followers were clearly more optimistic about the effectiveness of "Schumpeterian" competition—that competition for the market could replace competition in the market to ensure economic efficiency—than the analysis of chapter 5 would suggest. We have seen that dominant market power can persist and that the threat of competition may lead to neither high levels of innovation nor low levels of profits.

There are no general proofs of the efficiency of the market economy in producing innovation. And as we commented in chapter 1, the dominant paradigm—the competitive equilibrium model—does not even address the issue. The central theorem of welfare economics (usually called the *fundamental theorem of welfare economics*, Arrow [1951a] and Debreu [1959]) formalization of Adam Smith's is the notion that competitive market economies lead to [Pareto] efficiency through the invisible hand, *assumed* that technology was fixed, or at least *exogenous*, unaffected by anything that market participants might do. The followers of neither the Arrow-Debreu nor the Schumpeterian tradition have succeeded in remedying the obvious lacuna in their analyses: The former have not shown that competitive markets with innovation are Pareto efficient, and the latter have not been able to show that Schumpeterian competition would ensure economic efficiency *even in the production of innovations*. The reason that they failed is simple: There is, in fact, *no presumption that markets where innovation is endogenous are efficient*.

Until the development of the modern theory of the economics of information, the presumption in conventional economics was that markets, by themselves, result in efficiency (with well-known exceptions, such as those associated with pollution). Schumpeterian competition *seemed* to create a similar presumption for dynamic economies, in which the center of attention is on innovation.

As a result of our research on the economics of information over the past four decades, even as we began our research into the economics of learning, we were less sanguine about the outcome of these market processes. In our earlier work, we had established that markets in which information was endogenous (or risk markets were not perfect) were generally not efficient.[1] Knowledge (say, about new technology) can be viewed as a special kind of information, sharing the same essential properties (to be discussed). Not surprisingly, then, the *economics* of information and the economics of knowledge were very similar.[2] Given the similarities between "information" and "knowledge," it was clear that economies in which knowledge (including knowledge about technology) was endogenous would also be inefficient. Moreover, the fact that R & D was risky, and the risks could not be insured against, strengthened presumption *against* unfettered markets.

The market failures that we describe take on many forms: the structure of the economy, the allocation of resources to R & D—not only the total amounts but the allocation to various sectors and projects and the directions in which research is pursued—the levels of inputs (both in total and across sectors), the choice of technique, and how much information and knowledge gets disseminated. This chapter thus describes the numerous ways in which markets fail to allocate resources efficiently toward innovation and, more broadly, fail in creating (on their own) as dynamic a "learning" society as they might. It serves as a prelude to the rest of the book, where we describe how government actions can help create a learning economy. In the ensuing chapters, we will be able to pursue only a few of these distortions, explaining how government actions might partially correct them and help create a learning economy. But underlying all these market failures is a simple point: where there is learning (i.e., nearly always), there are marked divergences between social and private returns.

This chapter is divided into six sections. The first examines the distinctive properties of knowledge—why the production of knowledge is different from the production of steel and why, as a result, while there is some presumption that markets make the "correct" decisions about the level of production of steel (or would, in the absence of learning and other externalities), the same is not true for the production of knowledge. The second section delves further into several of the key market failures—why social and private returns to learning and R & D are likely to differ markedly.

In the third section, we ask whether innovation is always welfare enhancing. At a broad level, we know the answer. Many of the financial innovations directed at circumventing regulations intended to enhance macroeconomic stability contributed to the instability of the economy and played an important role in the 2008 crisis. We focus, however, on a narrower but long-standing question: Is it possible that the market focuses excessively on saving labor, contributing to higher levels of unemployment and more inequality?

The fourth section looks at the innovation process in broader, evolutionary terms. Though the results of our analysis differ markedly from that of the standard neoclassical analysis, the tools of analysis we use in most of this book are standard. But looking at the issues of innovation through the lens of evolutionary processes leaves us no more sanguine about the efficiency of market processes.

The fifth section takes a comparative institutional approach: Are there some forms of economic systems which are more conducive to learning and innovation than others?

In the final section, we make some more general observations about innovation and the nature of our society.

Before beginning the analysis, some further introductory remarks are in order. In comparing innovation under different market structures with that under the "socially optimal" arrangements, it is important to note two critical distinctions: First, while we saw that monopoly results in a lower level of output in the monopolized sector, and hence the benefit from innovation is lower than in the social optimum, the level of innovation, *given the level of output*, may still be optimal. We often speak of "constrained optimality." For example, given that the government cannot eliminate the monopoly, is the level of innovation "optimal"?

The analysis that follows (elaborated further in part 2 of this book) shows that, in general, markets are not even constrained optimal.

Second, societies with the highest levels of innovation (whether from learning by doing or from investment in research and development) may not have the highest levels of societal welfare, since there is a trade-off between current consumption (well-being) and future consumption. Indeed, later we shall note some instances where there is too much innovation, and especially too many resources may be allocated to innovation in a particular sector or in a particular direction. In chapter 5, we described how different market structures or institutional arrangements affect the pace of innovation, but one should not confuse such an analysis with an analysis of which kinds of market structures or institutional arrangements enhance societal welfare. We will show that, in general, there are welfare-enhancing interventions in the market.

Still, it is important to note that, while *ex ante* societal welfare may not be enhanced if there are *excessive* investments in innovation, in the long run, citizens in those countries which have engaged in these excessive investments are better off. They benefit from the sacrifices in consumption made by their forebearers.

1. Distinctive Properties of Knowledge

Knowledge is different from conventional goods in three ways that result in markets normally not being competitive (and, of course, if markets are not competitive, they typically won't be efficient) and, even if competitive, normally not being efficient.

Knowledge as a Public Good and Learning Externalities

Most importantly, knowledge is a Samuelsonian (1954) *public good* (Stiglitz 1987b, 1999a)—the marginal cost of another person or firm enjoying the benefit of knowledge (beyond the cost of transmission) is zero; usage is nonrivalrous. Moreover, it is typically difficult, if not impossible, to exclude others from enjoying the benefits of the knowledge produced. These were the spillovers that we emphasized in chapter 3, and these spillovers (externalities) play a central role in the analysis of this book.

Markets are not efficient in the production and distribution of public goods—including the production and dissemination of knowledge.

So long as there are spillovers, there will be underproduction—the firm won't take into account the benefits that accrue to others.[3] Even when an innovator becomes rich as a result of an innovation, what the innovator appropriates is often but a fraction of what the innovation has added to GDP. This is especially obvious in the case of those who have made the most important discoveries—those who have made major contributions to the advances of basic science and technology receive rewards that are substantially below their social contributions. Think of Alan Turing, James Watson and Francis Crick, Timothy Berners-Lee, or even the discoverers of the laser/maser and the transistor.[4] It is fortunate that most of these were not motivated by material rewards—a hint that the obsession with such rewards by the advocates of stronger intellectual property protection may be misguided (as we argued in chapter 3).

But externalities are more pervasive than these examples illustrate, and most innovations cannot be protected even by the most stringent intellectual property protection laws. Some knowledge has spillovers to particular industries, or to particular processes in particular industries, while the benefits of other knowledge can be more widespread, as we noted in our earlier discussions. The assembly line and then just-in-time production transformed production processes across wide swaths of the economy. The discovery of rayon showed that artificial fibers could be created and provided impetus for others to find alternative fibers not covered by the patent. Individuals who learn about better ways of doing business transmit that knowledge when they move from one firm to another.

There are two properties of public goods—nonrivalrousness and nonexcludability—and intellectual property aims to partially address the second. It attempts to "solve" the excludability problem, simply by not allowing others to use the knowledge for a limited period of time without the consent of the producer of the knowledge (the owner of the patent). In doing so, it attempts to reduce the extent of spillovers. But attempts to "capture" the returns to knowledge by restricting its dissemination introduce another distortion—in the efficient utilization of knowledge.

OPTIMAL RESOURCE ALLOCATIONS WITH LEARNING AND LEARNING EXTERNALITIES: BASIC INTUITION It is easy to describe the efficient resource allocations without learning: In each period, the marginal benefit of producing one more unit of a good must equal its marginal cost.

The value of the marginal product = marginal cost today. (1)

In the case of a good produced by labor alone, this can be put more formally: the marginal rate of substitution between the good and leisure (which should be the same for all persons) should equal the marginal rate of transformation, that is, the marginal product of labor.

Learning while producing or investing has future benefits—lower future production costs—and this needs to be taken into account. This can easily be done:

The value of the marginal product + *total* future cost savings
= marginal cost today. (2)

The competitive equilibrium with learning will entail a higher level of production, because of the benefits from learning. But with learning externalities, the social benefits of learning (the total future cost savings) are far greater than the benefits of learning that accrue to the firm.

Social value of learning > > private value of learning.

Hence, the level of production will be smaller than is socially optimal. In fact, with very small firms, the value of learning to each firm from its own production is very small, even if the value of learning to society as a whole (to all firms in the industry) may be large. The result is that the competitive equilibrium with learning and that without learning will differ little. It should be clear, then, that there will be too little production, and thus too little learning.

Imperfections of Competition

The second distinctive property of knowledge follows from the first: innovation is marked by returns to scale. As we discussed in chapter 5, in certain limiting cases, where there are no offsetting diseconomies of scale, there is a natural monopoly. In other cases, there may be an oligopoly. In all cases where R & D is important, competition is likely to be limited.

These results hold even if there are spillovers, so long as spillovers are imperfect. For as long as that is the case, the firm engaging in learning (R & D) will have a cost advantage over rivals that engage in less learning (R & D). In practice, there are likely to be some imperfect spillovers, so that markets where innovation is important are likely to be marked by large externalities *and* high levels of imperfections of competition.[5] Thus, both the level of production and the level of innovation will be distorted.

The only way that there can be effective competition in the situation just described (constant marginal costs of production) is that there be *full* spillovers to others in the same industry; but if that is the case, then each firm will free ride off the research efforts of others and, if the number of firms is large, there will be no incentive to engage in R & D. If knowledge (learning) is a by-product of production or investment, each will have insufficient incentives to produce or invest. Either way, the economy will not be efficient.

As we also noted in chapter 5, competition can be sustained if there is some offsetting force for decreasing returns. But there is a dilemma, the nature of which will be made clearer in this chapter: With more firms, the distortions in production will be reduced, but distortions in innovation may increase.

Not only are there likely to be multiple market failures, but imperfections in one arena are likely to lead to failures in others. As Arrow (1962a) pointed out fifty years ago, the production of knowledge is often a joint product with the production of goods (there is learning-by-doing), which means that the production of goods themselves will not in general be (intertemporally) efficient.

THE CASE OF MONOPOLY/MONOPOLISTIC COMPETITION

Monopoly (or more accurately, monopolistic competition, where there is a single firm producing any commodity, but there are many producers producing different products vying for the consumers' dollars) provides a limiting case. When competition is restricted, market allocations are not efficient. But now there are two inefficiencies: In addition to the static inefficiencies associated with the exercise of monopoly power, there may be dynamic inefficiencies.

As a first approximation, these inefficiencies are reflected in the condition for optimum production:

$$\text{Marginal revenue product + future cost savings } \textit{to the firm} \\ = \text{marginal cost today.} \tag{3}$$

Equation (3) should be contrasted with (1) and (2): the monopolistically competitive firms underestimate the static benefit of production, ignore learning benefits to other firms, and, because production may be lower, assign a lower value even to firm cost savings (than would be the case at the social optimum). Products in which firms have more monopoly power will have less production, and the lower production will lead to less learning. Productivity growth in these sectors may, accordingly, be slower.[6] In addition to the static consequence of the loss of consumer surplus from underproduction, there is a dynamic cost: The lower learning and higher costs in subsequent periods associated with monopoly today result in lower output in future periods.

Of course, labor not used in the monopolized sector gets displaced to other sectors, but if those sectors are sectors with less learning, the overall rate of growth of the economy is reduced. Moreover, monopoly power will result in lower real wages; lower real wages will normally result in lower equilibrium labor supply and, hence, less learning.

One of the important methodological implications of the analysis is that not only must one simultaneously consider market structure and innovation (both are endogenous), but the analysis must be conducted within a general equilibrium framework. In a partial equilibrium context, one might conclude—as Schumpeter did—that monopoly is better than competition because it internalizes the benefits of learning, without noting adverse general equilibrium effects, arising from impacts on the *pattern* of production and overall labor supply.

Understanding the structure of learning and knowledge dissemination is essential to understanding efficient production. We are concerned with societal learning, not just sectoral or firm learning. For example, some sectors may have stronger learning curves; that is, the elasticity of learning may be larger for a firm. But what matters is not just the ability of a firm or sector to learn but also the benefits that sector (firm) transmits to other sectors (firms) and the extent to which it does not appropriate for itself the benefits of the learning. If learning in one sector generates more externalities to other sectors than do others, production in that sector should be increased (relative to what it would be in the market equilibrium that ignored these learning externalities) at the expense of others. The dynamic (future) benefits need to be offset against the static (short-run) costs.

Imperfect Risk and Capital Markets

The third distinctive property of knowledge production leading to pervasive market failure in innovation—one which interacts with the previous two—arises from the fact that R & D is inherently risky, and risk markets with innovation are inherently imperfect. The outcomes of research and learning cannot typically be fully foreseen: Research is an exploration into the unknown. As research proceeds, new ideas are developed and new (and unanticipated) products may emerge.

One cannot buy insurance against the risk that a research venture will prove fruitless or that there will be little learning as one gains experience. Part of this is explained by theories of asymmetric information. There are inherent problems of moral hazard and adverse selection. The researcher is more likely to know more about the likelihood of success or failure of a research venture than any outsiders. Insurance markets where information asymmetries are large often don't exist, and even when they exist, the insurance offered is limited.

There are, in addition, fundamental conceptual issues: One cannot buy insurance (an Arrow-Debreu state contingent security) against an event (the discovery of nuclear power) prior to the conceptualization of that event. Nor can one have an insurance market for the "explosion" of a nuclear reactor before the development of nuclear power, and one cannot have an insurance market on nuclear power replacing fossil fuels before the development of the understandings of modern physics on which nuclear power was based.

The absence of adequate risk markets presents a barrier to entry: Large, well-capitalized firms are better able to bear the risks associated especially with large-scale investments in research. This reinforces the conclusion reached earlier that markets where R & D is important are likely to be marked by significant imperfections of competition.

The absence of imperfect risk markets compounds the problems posed by imperfect capital markets. The modern theory of information asymmetries has helped explain why capital markets are often highly imperfect: why, for instance, they may be marked by credit rationing. Capital market imperfections can be particularly adverse for learning. Because R & D investments (or, more generally, "learning investments"[7]) typically cannot be collateralized, unlike investments in buildings, machines, or inventories, it is more likely that there will be credit and equity rationing, leading to underinvestment in these areas, compared to others.[8] (Moreover, as we have noted, there are other fundamental reasons for capital market imperfections: a borrower with a good idea worries that telling a potential lender about the idea will lead to his stealing the idea, or in some other way taking advantage of that knowledge to advantage himself, at the expense of the borrower.)

Because of imperfections in capital and risk markets, firms act in a risk-averse manner, particularly in the presence of bankruptcy costs (Greenwald and Stiglitz 1993), and this discourages investment in riskier innovation. (It also explains why, as we argued in chapter 4, cyclical fluctuations are so bad for innovation: in economic downturns, investments in R & D are among the categories of expenditures that suffer the most.)

There are further distortions associated with imperfect risk and capital markets. Learning typically requires forgoing output or bearing risk today, in the hope of higher output in the future; hence even when firms can appropriate future benefits that derive from their research and learning today, they may discount those benefits with a high discount factor, resulting in suboptimal levels of learning and research.

Furthermore, the entry of one firm, and its investments and research strategy, has an effect on the riskiness of investments by others, which it obviously fails to take into account. But a change in risk affects their investment, especially because (for reasons already explained) firms act in a risk averse manner.

Given the pervasive market failures that we have already identified—the public good nature of knowledge, the pervasiveness of externalities/knowledge spillovers, the limitations of competition, the imperfections of risk and capital markets—there is no reason to be sanguine that markets are efficient in the level or direction of innovation. In fact, matters are worse: there are a number of other market failures that are intimately associated with innovation. Moreover, the various market failures are interlinked; one market failure can reinforce another. We explore these issues in the next section.

2. Further Reasons Why Markets for Innovation Are Inefficient

Markets fail to produce efficient outcomes whenever private rewards and social returns differ. This occurs when there are externalities or imperfect competition, imperfect risk markets, imperfect capital markets, or information asymmetries—and these "imperfections" are inherent and important in the innovation process itself. The previous section emphasized, for instance, that there are inevitably important spillovers, competition is necessarily imperfect, and innovation investments are risky, often requiring large up-front investments, so the absence of perfect capital and risk markets is consequential. This section highlights further limitations: private rewards differ—and sometimes exceed—social returns, and there are large coordination failures.

Private Rewards and Social Returns

Our earlier discussion highlighted that firms (individuals) appropriate less than the full value of their societal contributions from learning and R & D; there are important spillovers. This by itself might suggest that there is a presumption of too little learning or investment in R & D. But there are some circumstances in which private rewards may exceed the social returns, with the consequence that there may be excessive investment in R & D, and the problem may be exacerbated by (inappropriately designed) intellectual property regimes. In the paragraphs that follow, we illustrate several important instances in which this is likely to occur, especially in the context of a poorly designed patent system. The first five of these are examples of what we described in chapter 5 as "rent stealing," where part of the profits accruing to an innovating firm are profits that otherwise would have accrued to some other firm. Firms can garner profits for themselves not just by making consumers better off but by making their rivals worse off.

RENT APPROPRIATION AND ME-TOO INVENTIONS An obvious example is "me-too" innovations,[9] where researchers try to develop a product essentially identical to one already on the market. The object of the research is to find a way around the patent and to grab a share of the patent holder's profits (rents). While me-too innovations are particularly pronounced in pharmaceuticals, they arise in other sectors as well. This illustrates a general aspect of the returns to innovation: *the rents captured by a monopolist are not directly related to the increase in consumer welfare (surplus) associated with the innovation*. Some of the rents are rents *diverted* from other firms. Firm *i*'s profits from an innovation are *not* a good measure of the social contribution of its innovation, as they would be in a competitive market, where the price of the product was given and equal to the marginal cost of production, and the increase in profits would simply measure the reduction in societal resources required to produce the firm's output.

SOCIAL AND PRIVATE RETURNS TO WINNING A RACE Patent races provide another context in which private returns may exceed social returns. In an innovation process, the social return is only that the innovation is available earlier than it otherwise would have been. Myriad Genetics obtained a patent on the BRAC genes (which are critical in assessing the likelihood of getting breast cancer). The gene would have been discovered shortly later, as part of the more systematic attempt to decode the human genome. Thus, while Myriad has made large profits, its social returns were small. Indeed, arguably, because it has exercised its monopoly power, charging high prices for the tests to detect the presence of the gene and preventing follow-on research, including the development of superior tests, the social return has been negative—and depending on how one assesses the value of lives lost as a result of women who could not afford the test at Myriad's monopoly price, perhaps very negative. Had Myriad not entered the fray, the test would have been available at a very low, competitive price (see Stiglitz 2006a; Azvolinsky 2012; Goozner 2010). The distortion here can be viewed as another form of "rent-stealing"; by engaging in faster research, one appropriates the rents that would have occurred to others, to those who would otherwise have made the discovery.

It should be clear that the race to be first can even arise in the absence of a patent when there is a first mover advantage to the first entrant into a market because of, for instance, the development of brand loyalty or a "lock in" to his technology.[10] Under the patent system, the race to be first has further adverse consequences: the equilibrium may entail smaller innovation steps made more frequently than would be the case in an optimally designed innovation process.

MONOPOLY RENTS AND THE ENCLOSURE OF THE COMMONS Another reason that social returns may differ from private returns is that the profits of the winner of the patent typically include not just monopoly rents but, in some cases, a return to the privatization of knowledge that was previously in the public domain.[11] Moreover, because success in getting a patent converts what would be a public good into a private good, while success in challenging a patent converts what would have been a private good into a public good—that is, the opposition to the granting of a patent is itself a public good—there will be excesses in the granting of patents.

THE COMMON POOLS PROBLEM We noted earlier that a primary determinant of the pace of innovation is the pool of knowledge from which others can draw. Innovations both contribute to and (especially with the patent system) take away from the pool of publicly available ideas that can be drawn upon. In a private market economy, each firm not only has an incentive to take out of the pool of available knowledge as much as it can (by getting as broad a patent as possible), it also has an incentive to contribute as little to the pool of knowledge as it can—any knowledge that it has that others do not gives it a competitive advantage.

There is a correlate market distortion: a tendency in a market economy for excessive secrecy. Even though there is a social value in others having access to already produced knowledge, there are strong private incentives to restrict the flow of knowledge.[12]

The fact that there is a common pool of knowledge from which all can draw freely gives rise to what is called the common pool problem; there is (focusing on this effect alone) a tendency for *excessive* investment and excessive entry. In making their decisions about how much to invest, each ignores the adverse impact on others as it takes out of the knowledge pool. But the fact that there is excessive entry and investment does not mean that the pace of innovation is faster than in the social optimum. Because each firm has an incentive to take out of the pool of knowledge from which others can draw as much as it can and contribute as little as it can means that the size of the pool of knowledge available will be lower. But with a smaller pool of knowledge to draw upon, the pace of innovation will be slower.[13] Obviously, the design of the intellectual property regime affects the nature and extent of these market distortions. Stronger disclosure requirements (effectively enforced) will mitigate this market distortion. Tighter intellectual property regimes and broader patents may exacerbate it. Thus, even if a tighter intellectual property regime leads to more innovation *given a fixed pool of knowledge,* taking into account the adverse effect on the pool of knowledge, the pace of innovation may be lowered (see Stiglitz 2014a).

HOLDUPS "Holdup" patents provide another instance in which social returns are almost surely markedly lower than private returns, and they reflect another major source of distortion in market-driven innovation, related to bargaining problems. Modern inventions often require a number of ingredients (ideas), each of which can be patented separately. Thus, putting the product together requires agreement among a large number of patent holders. Some patents are less important than others. For example, the inventor can invent around the patent, though at some cost. There can be asymmetries of information. In the presence of such asymmetries, bargaining often leads to inefficient outcomes. Efficiency clearly requires the full utilization of the available knowledge, but as each side attempts to show its determination and to disguise the costs of a lack of agreement, bargaining sometimes breaks down. Firms are then forced to invent around the patent, which not only entails diverting scarce research dollars into duplicative research but may also result in costly delays in bringing the product to market.[14] Beyond this, there are often large wastes of resources in litigation expenses.

Holdup patents are used by patent trolls to extract rents out of successful innovators, claiming that they have infringed on their patents. Given the high costs of inventing around the patent or litigating, patent trolls can often extract handsome sums, reducing, by the same amount, the returns obtained by the "true" innovator. Thus, the returns on their "innovations" exceed the social returns, while those of the "true" innovators are less than their social contributions.[15]

EXTENDING MARKET POWER We noted earlier that markets where innovation is important are likely to be characterized by imperfect competition. Often, research is directed at extending (and increasing) market power—including deterring entry—rather than increasing the well-being of consumers. Pharmaceutical companies look for small innovations that enable them to "evergreen" their patents—giving their newly patented product a slight advantage over their old products as the patents on those products expire and they become produced by generic manufacturers.

CONSUMER SURPLUS In a standard competitive model, there is a close correspondence between an increase in profits and an increase in consumer welfare. But in models of innovation, a firm fails to capture the benefits of its learning and research because of spillovers;

in addition, there may be significant increases in consumer surplus. Each innovation builds on others. Learning today provides a higher base from which future learning starts. Future firms will accordingly face lower costs. Over time, prices of goods fall significantly, quality improves, and consumer surplus increases. In deciding the pace and direction of innovation, innovators do not take into account the effects on consumer surplus. For a large innovation, even in the short run, even a monopoly innovator does not capture societal benefits; there is still an increase in consumer surplus. The result is that the level of investment in and pace of innovation will be less than is socially optimal.

OTHER DISTORTIONS Whenever the price system "fails," market participants will have distorted incentives for innovation. Without a price for carbon, there are obviously no incentives to find innovations that reduce carbon emissions. It is thus not a surprise that our allegedly innovative economy has done so little to curb carbon emissions.

Our earlier work (1986) showed that market failures are pervasive in the economy whenever there are asymmetries of information or imperfect risk markets (that is, always), and a corollary of that result is that there will be market failures in the direction of innovation. The next section shows that there will normally be excessive incentives for labor augmenting technological progress, leading to higher levels of unemployment.

One way government responds to market failures, including externalities, is by adopting and implementing regulations. But that provides an incentive for the development of innovations that can circumvent the regulations, as illustrated by the financial sector, in the years prior to the 2008 crisis. In doing so, these innovations imposed high costs on the rest of society.

Coordination Failures

We champion the virtue of private markets in solving the complex coordination problems that are required for our large, interdependent economy to function. Prices play the central role in that coordination. Successful innovation, too, requires coordination, but prices don't (and can't) play the role that is usually hypothesized in the "normal" context of a market economy in the absence of innovation. In fact, matters are worse: Secrecy that is central to much of the market production of knowledge (part of the attempt to increase the degree of appropriability and to enhance the likelihood one will win the patent race) means that coordination is difficult.

There are many dimensions to socially desirable coordination. If research is uncertain, but additional research enterprises are imperfectly correlated with research enterprises already being undertaken (say, in producing a new product or reducing the cost of production of an existing product), then there is a social return to additional entry. We can easily describe optimum entry, where the marginal social value of an additional entrant equals the extra cost. But if no one knows who else is undertaking research, it is hard to achieve this. Moreover, the social optimum entails an optimal diversification of research projects, but again, with secrecy concerning what other researchers are doing, it is unlikely that this optimum will be obtained. There is likely to be excessively duplicative research.[16]

Moreover, the value of invention A may depend on the existence of a complementary invention B. Unless A knows that there is a high likelihood that B will be produced, A will have limited incentives, and similarly for B. Sometimes this coordination problem can be internalized: a large firm undertakes (or at least coordinates) the various parts of the research project. Indeed, the development of firms with these capabilities is one of the major advances of the twentieth century. Still, competencies and skills differ, and knowledge about competencies and skills is limited, so that a firm may not be able to bring under one roof (or, more broadly coordinate) those most likely to succeed in each of the parts of the research enterprise. If it turns out different "ingredients" are patented by different parties, a bargaining problem, with the associated inefficiencies and potentials for holdups, may well arise.[17]

Interactions among Market Failures

There are important interactions between traditional market failures, like imperfect competition, and those associated with learning. Problems of appropriability of returns, imperfections of capital markets, and the absence of good risk markets result in barriers to the entry of new firms (entrepreneurs) and the exploration of new products—including products or processes that might be particularly appropriate for a developing country. As we have explained, they give an advantage to large enterprises.

Theory of the Second Best—and the Financing of Innovation[18]

As we have seen, to Schumpeter, the fact that there is a distortion associated with monopolies/imperfect competition in innovation economies was not, by itself, of too much concern. After all, the fixed costs of financing R & D or learning had to be paid for somehow. Indeed, that was an implicit aspect of the argument that contestable markets (that is, markets where potential competition is so fierce that price is driven down to average costs) were efficient. Even if price equaled average cost (i.e., there were zero profits) rather than marginal costs, so that compared to the standard first-best resource allocation, there was a distortion, the fixed costs had to be financed *somehow*; and however the fixed costs were financed would impose a cost to society.

We can thus ask the question, What is the optimal way of financing the public good of research? Having it financed by a monopoly is *not* generally optimal. First, as we noted earlier, even with potential competition, monopoly profit—after paying for the cost of innovation—is not driven down to zero. Second, the incentive of monopolies is to increase profits in any way (legally) that they can, and that includes expending resources to reduce the elasticity of demand, which allows them to raise their price. Innovation too is directed wrongly—it is directed at strengthening and extending monopoly power and the profits derived from monopoly power; and those objectives are at odds with innovation directed at enhancing societal welfare.

Third, relying on patents and the monopoly profits to which they give rise to finance research results in an underutilization of knowledge. As we have noted, research is a fixed cost, and there is no marginal cost to the use of an idea, so that knowledge should be freely provided. But that would imply that the producer of information (knowledge) would receive no returns. Thus, it is inevitable that, in the absence of government finance, there be underproduction of knowledge (relative to the first best) and underutilization of the knowledge that is produced. The patent system (in principle) attempts to balance out the dynamic gains with the short-run costs of the underutilization of knowledge and imperfections of market competition, but it does so most imperfectly.[19]

On the other hand, when the government finances research and disseminates it freely, there is still a static distortion (from the distortionary imposition of taxes), but no distortion in the dissemination and use of knowledge. But a patent system can be viewed as financing the research by a tax on the buyers of the product with the innovation. Standard tax analysis would suggest that this "monopoly tax" is not the ideal way of raising the revenue. Such a tax does not minimize the distortions (dead-weight loss) associated with raising the requisite revenue. Also, the "monopoly tax" is a benefit tax, and while in certain circumstances one can argue for such a tax regime (those who benefit from the product pay for its development), in other cases, it is hard to justify. Someone suffering from a life-threatening disease is already unfortunate enough; to ask the patient, in addition, to pay an R & D tax to finance the development of the patient's own medicines is not consistent either with most ethical principles or with social welfare maximization.

3. Socially Unproductive Innovation: Is Innovation Always Welfare Enhancing?

This and the previous chapters have explained why it is that the allocation of resources to innovation is not likely to be socially optimal. There is no presumption that markets are efficient, either in the amount of research or direction of research and learning. We have emphasized, in particular, the failure of market participants to take into account externalities—the benefits that their learning has for others. The possibility that innovation may not be welfare enhancing was evident nowhere more than in the financial sector, where much of the innovation was directed at circumventing regulations that were designed to enhance the stability and efficiency of the financial system. The result was that, as Paul Volcker pointed out,[20] it was hard to identify innovations that had increased the productivity of the overall economy. The innovations had led not to better risk management and resource allocations but, rather, to more risk and a massive misallocation of capital.[21]

Historically, there have often been instances in which significant groups within societies have resisted innovation, most notably, the Luddites in the beginning of the nineteenth century, who saw modern machines as leading to unemployment and impoveishment. While increases in productivity *in principle* could make everyone better off—the production possibilities curve moves out—in practice there are always winners and losers. Innovations that reduce the demand for unskilled workers decrease their wages, even if it increases the wages of skilled workers. The statement that such skill-biased innovation could be welfare enhancing is usually taken to mean that the gains of the skilled workers are more than sufficient to compensate the losses of the unskilled workers. But while the skilled workers *could* compensate the unskilled workers, such compensation seldom occurs. Thus, there are winners and losers. And if, as has been happening in the United States and many other advanced industrial countries, the losers are those at the bottom of the income distribution, then innovation can contribute to growing inequality. In this situation, whether societal welfare is increased depends on how one weighs the benefits to the relatively rich against the losses to the relatively poor.[22]

More recently, however, we (with several coauthors—Delli Gatti et al. 2012, 2013) have shown that with market imperfections and societal rigidities, all (or at least most) groups in society can be worse off. In the 1920s, productivity increases in agriculture were so large that (especially given the inelasticity of demand for agricultural goods) incomes in that sector declined. With perfect mobility, the surplus agricultural workers would have moved into the urban sector. But there are significant costs to the mobility of labor, and with wages in agriculture declining and the value of rural assets (like houses) declining as well, many in that sector couldn't afford to move to the city and obtain the skills that would make them productive there. Worse still, neither they nor the banks that provided credit anticipated these events. Hence, as incomes in the rural sector collapsed, those in that sector were left with a legacy of debt burdens, and banks faced massive losses. The result was a marked decline in demand for urban goods—so great that incomes in the urban sector itself fell. Innovation may have helped precipitate the Great Depression.

We have argued, by the same token, that improvements in productivity in manufacturing, leading to decreased employment and wages in that sector, have contributed to the current economic slowdown. Innovation requires economic restructuring, and markets often do not manage such restructurings well. But as firms make decisions that affect the pace and direction of innovation, they do not take these general equilibrium effects into account. Each small firm takes the course of wages and unemployment, for instance, as given; but collectively, as they make their innovation decisions, they affect the evolution of wages and unemployment. *Unfettered and undirected markets may result in patterns of learning and innovation that result in more inequality and higher* *unemployment than is socially desirable. There are other patterns that would enhance societal well-being*.

To see this most simply, think of there being a limited amount of resources available (to society, to the firm) for innovative activity. The firm can allocate these scarce resources, say, between innovations that save on natural resources (e.g., reducing the carbon footprint) and those that save on labor. But given that there is no carbon price, there is no incentive to reduce the carbon footprint. Even if there is unemployment, and there is a significant societal cost to increasing the number of unemployed, there is a private return to reducing labor inputs.

The same reasoning applies in cases where the mispricing of resources is less obvious. Economists have puzzled about the persistence of unemployment, even in countries without minimum wages or with weak unions. Markets set wages at levels above that at which demand equals supply. The theory of efficiency wages provides at least part of the explanation (Stiglitz 1974b; Shapiro and Stiglitz 1984): an increase in wages increases profits, either as a result of lowering labor turnover, attracting a more productive labor force, or inducing the current workers to work harder—not to shirk. Under these circumstances, the actions of a firm have an externality effect on others, which it does not take into account, but which matter.

Consider the Shapiro-Stiglitz model in which at any level of unemployment, there is a critical wage, below which workers will shirk.[23] If each firm innovates in a way which reduces its demand for labor at any given wage, the equilibrium level of wages falls and the equilibrium level of unemployment increases. There is a social cost to this increased unemployment, but no firm, in making its decisions about the direction of innovation, takes this into account. Thus, if firms have a choice between innovations which are more labor augmenting (i.e., increase the productivity of each worker, so that with the new technology, each worker is equivalent to, say, two workers under the old technology)[24] and innovations that are more capital augmenting, the firm's decision as to the direction of innovation depends on the relative shares—it will choose more labor augmenting innovations if the share of labor is high—and the market equilibrium will entail innovations which are excessively labor augmenting.[25] So too, if these efficiency labor effects are more important for unskilled labor than they are for skilled labor, and firms have a choice between innovations which are more skilled labor augmenting or more unskilled labor augmenting, it will choose innovations which reduce the demand for unskilled labor excessively (innovation is excessively *skilled biased*).

The efficiency wage model is the simplest within which to explore market distortions in the pattern of innovation, but similar results emerge in other models with (endogenous) market imperfections. Economic historians, such as Salter (1966) and Habakkuk (1962) have emphasized the role of "labor scarcity" as an inducement to labor-saving innovation. (Such explanations seem to have particular relevance in particular historical periods, e.g., in the period of the rapid expansion of the United States in the nineteenth century.) Standard economic theory has had a hard time understanding what labor scarcity might mean, other than a high price (or share) of labor.[26] But in models with costly information and highly differentiated labor, there is a natural interpretation: It may take time and resources to recruit a new worker to replace a worker that leaves. Labor-augmenting technological progress reduces not only the direct labor costs but these indirect turnover (search and recruitment) costs. But in economies with costly search, a decision by a firm to engage in more labor-augmenting technological progress—and thus in less recruitment—imposes externalities on other market participants, both on workers (who must now search longer to find a job) and on other firms (who now may face lower recruitment costs). Again, there is no presumption that the market equilibrium factor bias will be efficient; indeed, there is a presumption that it will not be (see Greenwald and Stiglitz 1988; Arnott and Stiglitz 1985).

More generally, we note that from the perspective of the firm, what matters is not just the wage or interest rate, as it might show up in the system of national income accounts, but the *effective* total labor and capital costs, which can differ markedly from the recorded labor and capital shares, for several reasons. First, because of taxes and fringe benefits, the cost of labor to the firm may exceed the wage that workers receive by a considerable margin. Second, if there is credit rationing, the "shadow" cost of capital may well exceed the interest rate charged; and if firms can't instantaneously hire workers of the particular type in which they are interested (there is, in this sense, a labor scarcity), then the shadow cost of labor will exceed the wage. Even a relatively small gap in time in being able to fill a position may be costly. By the same token, if machines are not fully reliable and cannot be easily replaced, a breakdown of a machine can be costly. Third, workers have to be managed. Strikes are costly. All of this requires scarce managerial time.

When a firm assesses whether to save on labor or capital, all of these costs are relevant.

Policy Implications

There are several important policy implications.

First, wage subsidies reduce the cost of labor, and it is the high cost of labor that induces firms to shift the direction of technological development toward excessive labor-saving and capital-using technologies. By the same token, when the Fed lowers the cost of capital dramatically (as it attempted to do after the Great Recession), it encourages labor-saving innovation. Thus, we observe the curious phenomenon of firms replacing unskilled labor (with presumably a low shadow price, given the high unemployment rate among unskilled workers) such as check-out clerks with machines, e.g., automatic tellers. While there are almost surely positive social benefits from the induced employment resulting from the increased aggregate demand from such investments, those benefits have to be set against the social costs of higher unemployment in the medium term as a result of the labor-saving innovation induced by the lower cost of capital. A full analysis of the intertemporal trade-offs would take us beyond the confines of this discussion.

Second, increasing the price paid by firms for environmental impacts (e.g., carbon emissions) shifts innovation away from labor-saving (-augmenting) innovation, again with positive effects on the distribution of income and employment.

Toward a More General Theory

By bringing together a plausible theory of wage determination with the theory of induced innovation, we have provided a general theory of growth and employment which makes sense of discussions of technological unemployment or job shortages—concepts that have no meaning in Solow's formulation. In this theory the distribution of income matters; it affects technology and the dynamics of the economy, and these in turn affect the distribution of income at later dates.[27]

Recent discussions of persistent unemployment and growing inequality have centered around labor-saving innovations and in particular on skill-biased innovation (Autor and Dorn 2013). Critics of such innovation are sometimes referred to as modern-day Luddites, and defenders of the market have claimed that one should not interfere with market processes; in the long run, they argue, everyone will be better off. Our analysis has suggested that such views may be Panglossian. Not only within their life span may workers *not* be better off—the benefits of the improvements may not trickle down—but additionally, the changes in factor demands may actually lead them to be worse off even in the longer run.[28]

We have shown not only that innovations may fail to improve the welfare of all groups in society—they may not result in Pareto improvement—but also that the outcome of market processes may lead to patterns of innovation that are not even output maximizing—they would not be Pareto efficient even if redistributions could be made costlessly. Indeed, there is a presumption that unfettered markets will not be efficient in the choice of factor bias and will lead to excessively high levels of unemployment.

4. Evolutionary Processes

The central message of this chapter is that in an innovation economy, there are marked discrepancies between social returns and private rewards, so that there is no presumption that markets yield efficient outcomes. To the contrary, the presumption is that they do not and that there is a role for government to "correct" the market failures.

The fact that private and social profitability may differ markedly also helps explain why naïve arguments about the positive benefits of evolutionary processes are wrong. These arguments are often invoked by those who believe in the market but understand that the standard (Arrow-Debreu competitive) analysis fails to establish the efficiency of markets.

The recent crisis has cast further doubt on the validity of these perspectives.[29] For instance, financial institutions that had understood better the nature of risk and undertaken more prudent actions (e.g., not undertaken excessive leverage) did not survive. Investors observed their seemingly lower returns and demanded that management be replaced. This is not just a hypothetical possibility; it actually occurred. To be sure, those who argued for greater prudence can say, "I told you so." But firms (and their management) that were wiped out in the "creative destruction" of the process of irrational optimism and deficient risk analysis are not easily brought back to life.

The critique of the standard argument of evolutionary selection is that it makes *both* type I and type II errors. Firms and individuals that did well, and survived at least for a long time, were not necessarily those that contributed the most to societal well-being or even had attributes that suited them for long-run survival; rather, they were the firms that were well suited to take advantage of the irrational exuberance and the potential for exploiting the poor and market irrationalities that the era of deregulation had opened up. And those firms that were eliminated were not necessarily those that should have been.[30, 31]

Reward structures have allowed those who led the economy to the abyss to walk away with billions—less than they would have had if their flawed analyses had been right, but far more than they deserve, given the costs that they have imposed on the rest of society. With their wealth accumulation, they can exercise undue influence on the allocation of societal resources for years to come.[32]

Four critical insights help explain why evolutionary processes may not be efficient. The first, and most basic, is this: A necessary condition for evolutionary selection processes to work well is that profits are a good measure of social contribution. If that were the case, the firms that survived—had high profits—would be the ones that were making the most important social contributions; the firms that were making losses would be those making a negative social contribution, using up more resources than the value they created. But a central message of this book is that, particularly in the arena of innovation, profits may be a particularly bad measure of social contribution. More generally, evolutionary processes fail to produce efficient outcomes precisely in the same circumstances in which markets traditionally "fail," i.e., fail to produce efficient outcomes.

Second, markets are myopic. They ascertain how well firms are doing today—though because of accounting problems, so evident in the scandals that marked the beginning of this century, they perform this task very imperfectly (see Stiglitz 2003). They have a hard time ascertaining who will do better over the long term.

Moreover, even if a firm might do well over the long term, capital-market imperfections may mean that it will not be able to get the funds to survive now, if it is losing money. Thus, firms that may be more "flexible" and adaptable for changing circumstances might do well in the long run but not survive in the heat of the short-run competition. There may be firms that are better suited to the current circumstances.

They may compete intensely enough that the more adaptable firm has losses and can't survive.

But matters are even worse. Firms that are irrational can exert a negative externality on others—and cause them not to survive. The standard argument in economics is that if a firm is irrational (say, has irrational exuberance about the future of housing prices), it will pay the price—it will *eventually* lose on its speculation. But in its irrational exuberance, it can bid capital away from more rational firms, forcing them to pay a return beyond a level which they can sustain. This is not so much true if there is a single such irrationally exuberant firm, but will be especially true if there are many such firms—as was the case prior to the breaking of the bubble in 2006/2007.

The final important idea is the irreversibility of death. We have stressed that firms embody institutional knowledge—knowledge that is more than (and different from) the knowledge that is embedded in each of the individuals who is part of the organization. When an institution dies, much of the embedded knowledge disappears with it. And once a firm dies, when circumstances change—including circumstances under which it would have flourished—it does not come back to life. A new firm bearing some resemblance to the old might be created, but that entails large investments, sunk costs which might not be undertaken unless the expected returns are quite high.

Schumpeter stressed the importance of *creative destruction*. Firms with deep pockets and irrational exuberance may enter and drive out incumbents who are more rational and who are, in fact, better suited for long-run survival. That the former firms eventually die too is little comfort for those who disappear. This and previous financial crises illustrate that the externalities arising from such irrationalities may be economy-wide: the credit bubble imposed large costs, as we noted, on the rest of society.

5. Innovative Economic Systems[33]

This entire book is predicated on the notion that the level of innovation in a society is a function not just of a single policy (such as the enforcement of anti-trust policies, that might affect the level of competition, or the tightness of the intellectual property regime) but of the entire economic and social system. That was why we titled our book *Creating a Learning*

Society. It is natural, then, to ask if there are there some kinds of economic/social *systems* that are more conducive to learning. This is an exercise in *comparative economic systems,* a subject which was fashionable in earlier decades, focusing on the comparison between capitalist and socialist economic models. Here, our focus is more narrow: on the consequences of alternative economic and social systems for the level of innovation, and the comparison is among different versions of market economies.

In particular, we can ask what kinds of policies and institutional arrangements—what kind of economic systems—are most conducive to being an innovation leader—not just obtaining patents, but designing an innovation system that generates large and persistent increases in standards of living? Is it cutthroat competition? Or is the more gentle Nordic model, in which government takes on a larger role and in which a broad array of policies provide social protection and result in less inequality, more conducive to innovation?

We can also ask if we should expect that the policies of the follower differ from those of the leader, and if so, in what ways? Can we explain the successes of the Nordic model as a result of its policies being well adapted for the leader, or for the follower?

Acemoglu, Robinson, and Verdier (2012) have recently put forward the hypothesis that the Nordic welfare model may be all well and good for the follower, but the American style of cutthroat capitalism, with its high level of inequality and strong incentives, is better suited for the countries at the frontier. While contentions of such a broad sweep are hard to evaluate with any precision, similar sentiments have played a central role in policy debates and therefore it is important to assess them, marshaling whatever theoretical, empirical, and historical arguments can be brought to bear on the issue.

The analysis of this and the previous chapter should have made it clear that there is no presumption that unfettered markets will lead to the optimal rate and direction of innovation; that is, there is a presumption in favor of *some* government intervention. We explain here why the interventions associated with the Nordic model may in fact be highly beneficial to innovation.

Is the United States the Innovation Leader?

First, we need to dispose of one of the assertions of Acemoglu et al. (2012): that the United States is the "innovation leader." Assessing the level of innovativeness of an economy is no easy matter. Earlier, we explained that higher levels of *investment* in innovation do not necessarily lead to an enhanced pace of increases of standard of living overall, given the marked discrepancies between social and private returns. Even assessing the importance of any particular innovation may be difficult. Moreover, in a world in which knowledge flows in all directions, assessing the *origins* of any idea is nearly impossible. For instance, many of America's recent advances in medicine build on work done in the United Kingdom by Watson and Crick leading to the discovery of DNA. America's development of the computer rested on fundamental work done by Alan Turing in the United Kingdom. Parsing out the source of the "real" innovations is difficult, if not impossible.

The Swedish innovation of worker quality circles or the Japanese innovation of just-in-time production—neither of which were patented—may have had more profound impacts on American productivity than did multiple patented innovations. To be sure, Scandinavia benefited from Intel's innovations in chips, but presumably the value of those patented innovations would be (largely) captured in the profits of the patenting company and in the GDP of the originating country.

Interestingly, while many suggest that the United States has been highly innovative, say in the last thirty-odd years, it doesn't seem to show in GDP statistics, where increases in GDP per capita, or even estimates of total factor productivity growth, seem to be far lower than in the decades after World War II. There are several possible explanations for this. Perhaps GDP does not really capture the improvements in living standards that computer-age innovation is engendering. This may be partly due to the fact that GDP does not provide a good measure of well-being (see Stiglitz, Sen, and Fitoussi 2010), though there are reasons to believe that when full account is taken of, for instance, the increase in insecurity, economic performance is even more dismal than GDP statistics suggest.

Alternatively, it may be that as exciting as recent innovations seem, they are less significant than the enthusiasts believe. The United States may have made great strides in inventing better ways of targeting advertising, or designing financial products that are better at exploiting uninformed individuals. It takes innovativeness to design better ways to exploit and leverage market power, and this is likely to show up in higher profitability. But these "innovations" may not show up in GDP statistics.

Similarly, claims that the United States is more innovative than, say, the Nordic countries are not unambiguously supported by the data. By most accounts, Sweden and Norway have a higher standard of living or welfare (e.g., reflected in median income or UNDP's Human Development Index).[34] Moreover, output per worker hour in several countries exceeded that in the United States (Norway by 41%, Ireland by 15%, Luxembourg by 30%, Belgium by .5%), and in several (Germany, France, Netherlands, Denmark) the differences were small.[35, 36]

Acemoglu et al. (2012) put a great deal of emphasis on the high level of important patents in the United States. Putting aside technical issues such as differences in demographics and the presumed lower overall transactions cost associated with an American registering a patent in the United States versus a foreigner registering in the United States, there is a more fundamental issue: patents play markedly different roles in different sectors. In some sectors, like hi-tech and pharmaceuticals, they play a very important role, though in the former often more in a "defensive" way, to put oneself in a position to countersue when someone sues. In other sectors, like metallurgy, they play a very unimportant role.

By the same token, the number of citations is not necessarily a good index of importance. We referred earlier to two critical innovations—just-in-time production and quality circles. These were not patented, and, accordingly, there is no index of the number of citations. But there is little doubt of the profound effects. Or take another Swedish innovation: dental implants. Whether the original research spawned a large amount of follow-on research, with many citations, is not the critical determinant of the impact that this innovation had on the quality of life of hundreds of millions of individuals.

Also, the most important innovations, generating the most cited research, typically cannot be patented—from the Turing machine, to the discovery of DNA and electromagnetic fields.[37]

In short, the fact that the United States has a higher level of patents does not necessarily mean that the United States is more innovative. The United States may have focused its innovative efforts in those sectors where patents are important and where rent-seeking is encouraged, like the financial sector. If this is the case, then from the perspective of global innovation, it may be advantageous to have an ecology in which there are different institutional arrangements and no dominant one.

What Makes for an Innovation Leader?

Even if the United States were the most innovative country, it is hard to ascribe its position solely, or even mainly, to cutthroat competition. There are, in fact, multiple institutional and cultural factors that influence the ability of a technological leader to maintain that leadership position.

On the positive side, for instance, America's attitude toward bankruptcy (its acceptance of bankruptcy as part of the price to be paid for risk-taking in an innovative context) and development of the venture capital industry are two institutional characteristics that are highly conducive to innovation.[38] But even in these areas of strength, there are issues: U.S. bankruptcy law gives first claim to derivatives, and student loans can almost never be discharged, even in bankruptcy. This distorts the allocation of resources toward finance and away from higher education—distortions that almost surely result in less *real* innovation than there would otherwise be.

While the quality of its elite universities is clearly a favorable factor, the unevenness of the quality of its education—and the evidence deficiencies in average performance (e.g., as measured in PISA scores[39]) work in the opposite direction. So too does the fact that such a large fraction of its innovative talent has been diverted to finance (and zero sum activities *within* finance) and other rent-seeking activities.

While large corporations may have access to the extensive resources needed to undertake large, long-term research projects, the misalignment of the interests between management and shareholders may enervate innovation, especially of the kind that enhances standards of living. Particularly problematic are deficiencies in corporate governance (leading often to excessive short termism) and the bureaucratic processes that many large corporations have established as part of their control mechanism.

The success of the United States may have more to do with the large role played by the government than with the entrepreneurial role of the private sector (Mazzucato 2013). Even when we turn to private sector innovation, we find a picture quite different from that painted by Acemoglu et al. (2012) Probably the most innovative American firm during the twentieth century was a regulated monopoly, largely shielded from competition, with a research budget funded by, in effect, a tax on telephone service.[40] There are several reasons why that was so. Some of these are related to the fact that because a monopolist has a larger output (than, say, a duopolist, where total output is higher, but the amount produced by each firm is smaller) it has more incentive to bring down costs. Moreover, shielded from cutthroat competition, it could focus on the long run, including the benefits which it might receive from investments in basic research.

In chapter 5, we provided a more general analysis of the relationship between innovation and competition (however assessed), showing that it depends on a variety of characteristics, including the stochastic process of innovation, the substitutability among goods, the nature of the market barriers, etc. This suggests that the American model may be good for innovation in certain areas but adverse in others.

Finally, we note that particular historical circumstances played a central role in helping the United States attain the position in innovation that it holds today. The United States was not always the leader; in the nineteenth century, it borrowed voraciously from Europe. (See e.g., Chang 2001, 2002). Interestingly, even then, when it was a follower, it had a form of capitalism that was marked by high inequality—the extremes of the Gilded Age have only been reached in the Roaring 20s and in the first decades of this century. The innovator of the period, Germany, was the first country to introduce social security. The pattern seems to be the clear opposite of that suggested by Acemoglu et al. (2012)

World War II marked a turning point in U.S. technological leadership—a historical accident—partially, at least, a "gift" of that war—as large numbers of those on the forefront of science and technology fled to the United States.

This leadership was then reinforced as a result of government actions in response to the Cold War, which led to heavy investments in military research, which had large spillovers to the civilian sector (including, arguably, the development of the Internet). The large technological leadership of American universities, reinforced by World War II and government Cold War investments in the decades following the war, attracted some of the most talented young people from around the world, many of whom stayed in the United States.[41]

Why Might the Nordic Model Be Conducive to Innovation?

There are several key aspects of the Nordic model that may be particularly conducive to innovation. Earlier, we noted the importance of the (inherent) absence of a full set of risk and capital markets, both for the efficiency of the economy in general and for innovation in particular. Research is risky, and better systems of social protection can thus be more conducive to individuals undertaking research. Even high taxes can be conducive to risk taking; the government can be seen as a silent partner, sharing in the gains as well as losses, with the result that there will be more risk taking.[42, 43]

A major input into research is high-quality research personnel. Without government intervention, because of imperfections in risk and capital markets,[44] there will be insufficient investments in education. In the United States, with heavier reliance on private financing of higher education, adverse bankruptcy laws (in which student loans are essentially impossible to discharge), and the virtual absence of income-contingent loans, investments in education—especially in areas where returns are risky and limited, such as in science—will be more limited. And access to quality education by those whose parents have limited income will be greatly circumscribed.

Worse still, given the high cost of higher education and the skewed material rewards system, it is not a surprise that a disproportionate share of the most talented individuals have, in recent years, gone into finance; and while that may have resulted in a higher level of innovation in the financial sector, it has not resulted in a higher overall pace of innovation in the relevant sense (an increase in standards of living, or the pace by which standards of living increase). Indeed, much of the innovation was directed at figuring out better ways of manipulating the market, exploiting more those who were financially unsophisticated, enhancing the ability to leverage market power, and circumventing regulations that attempted to stabilize financial markets and reduce the risk of large adverse externalities.[45] While these innovations may have generated more rents for those in the financial sector, there is no evidence that they improved the overall performance of the economy.

Education is not the only critical factor that is complementary to private investments in innovation. Good investments in infrastructure can increase the returns to private investments (Field 2011) in general, including investments in innovation.

The Nordic model—with heavier public investments in education, technology, and infrastructure; progressive taxation that reduces incentives for rent seeking; and better systems of social protection—increases the willingness and ability for innovative risk taking. For an excellent discussion arguing that that is in fact the case, see Barth, Moene, and Willumsen, 2014. They go further, showing that in a vintage model of innovation, wage compression induces older vintages to be scrapped earlier, thus accelerating the process of creative destruction.[46] Moreover, they show how government policies can ensure that society as a whole benefits from innovations through, for instance, the active labor market policies and Keynesian demand policies that are part of the "Nordic model." Further, the Nordic model can lead to faster dissemination of ideas throughout the economy, which we have argued in earlier chapters is important for enhancing societal productivity (in ruthless competition, firms strive to keep whatever knowledge they acquire to themselves).

There are many more specific policies in the Nordic model that enhance innovation. Consider this question: Could innovation be encouraged by increasing taxes for financial and land speculation and using the proceeds to attract more into innovative activities by investing in science and technology education or paying scientists more? Standard arguments would suggest that higher taxes on land will not affect the land supply. And given the evident low (negative) marginal social returns to innovations in the financial sector, the reallocation of resources in ways that are associated with the Nordic model would presumably be "real" innovation enhancing. Or consider this slightly broader question: Could innovation be enhanced by taxing those at the top at higher rates and using the proceeds in a similar way? It has been argued that because much of the income is derived from rent seeking,[47] an increase in taxes at the very top has little effect on growth (Piketty, Saez, and Stantcheva 2011).

The Nordic model consists of exactly the kind of policies that one would expect to see in a leader. While it may not be optimal for all countries to follow the same model, those countries that aspire to be on the frontier should at least consider emulating some aspects of the model that has worked so well in the Nordic countries to maintain a high rate of growth not only in productivity but in standards of living as well.

Political and Economic Equilibria

The discussion so far has explored the consequences of alternative economic policies, but as is now widely recognized, public policies are enacted through political processes, which themselves are affected by the economy, including by the extent of inequality. We have to view the economic and political equilibrium as being jointly determined.[48]

It is easy to show that there can be multiple equilibria. In particular, there can be an equilibrium with a high level of inequality supporting low levels of public investments (including in education and technology), low levels of tax progressivity, and high levels of rent seeking, generating high levels of inequality. There can also be equilibrium with a low level of inequality with high levels of public investment, high levels of progressivity, a strong welfare state, and strong policies against rent seeking (the Nordic model). The representative individual is likely to be better-off in the latter—and so is the pace of innovation.[49]

There is no reason to believe that the United States has adopted the policies that it has because they are designed to maximize innovation, let alone societal welfare. Rather, the policies are simply the outcome of political processes in which those with money have disproportionate influence, an outcome that one might expect given its high level of economic inequality.[50]

This analysis suggests that the United States could increase the pace of innovation (and the level of economic welfare) by making some moves in the direction of the Nordic model. Not only would the institutional and policy reforms promote greater innovation directly, but by reducing inequality and the insecurity associated with innovation and openness, the reforms would generate more support for innovative policies and ensure that those displaced by innovation are "recycled"—retrained so they can be more productive members of the economy.

Many aspects of the Nordic model were explicitly designed with this political-economic equilibrium in mind. (See Barth, Finseraas, and Moene, 2012). The Scandinavian countries are small. To be prosperous, they have to be open to the outside world. But openness imposes high costs on many individuals—so too for innovation. And in truly democratic societies, if a majority of citizens are losers—even if a minority are large "gainers"—it will be hard to sustain policies supporting innovation and openness.

To sustain innovation and openness, one has to either move away from democracy (e.g., by moving toward a system where money has *more* influence), so that the winners have a disproportionate role in determining outcomes, or ensure that a majority of citizens are in fact better off—and that is the intention of the Nordic model.

Leaders and Followers

Not all countries can be leaders. Indeed, the aspiration of many developing countries is simply to close the gap that separates them from more developed countries. The policies that promote the follower's learning may be different from those that are optimal for a leader; for instance, the intellectual property regimes that are appropriate for each are likely to differ markedly. Thus rather than beginning from the presumption that there is a single economic system that is best for all countries, we should recognize that different countries may be in markedly different situations; there are leaders and followers. What is optimal for a leader, trying to move the frontier ahead (and perhaps maximizing innovation rents), may be different from that which is optimal for the follower, trying to take advantage of knowledge produced by others—trying to catch up, or at least not fall behind.

But, still, the economic framework that we have described (the Nordic framework) is more likely to be conducive to catching up than that associated with ruthless capitalism. We emphasize here, in particular, the importance of public investments in education and technology and in government policies that facilitate learning—including trade interventions and industrial policies that will be discussed extensively later.

Even countries that could *fully* close the knowledge gap may choose not to do so. There is a cost to doing so, and the cost is sufficiently high that a country may *choose* to remain a laggard.[51] There may then exist an international equilibrium, in which there are leaders and followers.[52]

Overview

Market failures affect both supplies and prices of inputs into innovation as well as the risk-adjusted private and social returns. The Nordic model can be thought of as addressing these market failures in a fairly comprehensive way. Policies affecting societal well-being directly—education, social protection (especially of children), unions, public investments in technology and infrastructure, active labor market policies, industrial policies—also affect the pace of innovation. Though there are some features of the American form of capitalism that are conducive to innovation, there are others that are not, and while there may be questions about precisely how strong its economic performance has been—say, in comparison to the Nordic countries—it is clear that what success it has attained can only be partially attributed to its markets and its form of cutthroat capitalism. Some is a result of a historical accident. Some can be attributed to its not-for-profit universities. Some can be attributed to strong government support.

There is an important role for government to play in both the leader and the follower countries: designing policies that lead to more innovation, ensuring that the knowledge gap doesn't increase, and ensuring that innovations disseminate well within the economy and that most citizens benefit. As we noted, without the appropriate frameworks, innovations can lead most citizens to be worse off, even if the winners could have compensated the losers.

In democracies, whether governments adopt policies that facilitate innovation will depend on the consequences of innovation for most citizens. The Nordic model, by ensuring that more of its citizens benefit from innovation and growth, has created a virtuous circle: a political regime that supports policies that facilitate innovation and ensures that the benefits of the resulting growth are widely shared.[53]

6. Broader Considerations: Innovation and the Nature of Society

Innovation shapes and is shaped by our society. Decentralized market processes typically pay little attention—for good or for evil—to these consequences. We have already noted two aspects: the effects on unemployment and the distribution of income.

The analysis so far has embedded innovation in a market economy in which, while there may be limited competition in the product market, there is a perfectly competitive labor market. But mobility is limited, and labor markets are often far from competitive. Employers—who manage the innovative process—may have an incentive to manage it in a way that enhances their bargaining process vis-à-vis workers, for that will lead to lower wages. Labor-augmenting innovations, which increase unemployment, do so. So do innovations which make workers more substitutable for each other. The pattern of innovation that has occurred in recent years, leading to lower wages for most workers, may not be just an accident of nature, nor may it even be the result of normal competitive market forces working themselves out; it may be the result of employers deliberately shaping the innovative process in ways which enhance their well-being at the expense of workers.[54]

Workers and management care not just about wages but about "control." Management, for instance, might like to reduce the scope for "agency" problems, where workers shirk or take actions which are not in the company's interests.[55] They may seek to reduce the scope for discretion. Innovations that increase the ability of management to monitor and control (the assembly line, just-in-time inventory systems) may be viewed as desirable by management, even when they are viewed adversely by workers.[56] There may be broader consequences to society of such changes in the workplace.

Of course, as firms engage in labor-saving (-augmenting) innovations, they "learn to learn"; they become better at this form of innovation.[57] This reinforces the process of labor-augmenting (-saving) innovation.

On the other side of the ledger, consumers, workers, and management may all get direct pleasure out of living in a more dynamic economy and society, and from the enjoyment of new experiences that results. These benefits too may not be adequately reflected in market prices and incentives.[58]

7. Concluding Comments

This chapter explains why the production of knowledge—or learning more generally—is different from the production of steel or other conventional commodities. While research over the past forty years has called into question the presumption that markets are efficient, in the case of a learning economy, the presumption is clear: It is unlikely that markets are actually efficient. In an innovation economy, Adam Smith's invisible hand is invisible because it's simply not there.

This chapter provided a list of key attributes of learning and innovation (characterized by fixed, sunk costs, being a public good, spillovers, etc.)—attributes which differ from ordinary commodities. This list explains the pervasiveness and importance of market failures associated with learning, and why the level, direction, and form of investments in learning in unfettered markets are not likely to be optimal. The existence of spillovers, for instance, means that those engaged in learning cannot appropriate for themselves the full social benefits of the learning, both today and in the future. This chapter has described a number of other reasons why the level and direction of innovation in a market economy may be far from optimal. Our list of determinants of learning also includes key societal attributes, like stability, which themselves are the result of public policy; markets by themselves do not necessarily result in the optimal level of macroeconomic stability. We should perhaps more accurately say that markets never exist in a vacuum. Society has to set the rules and regulations that govern them, like what kinds of contracts can be enforced and how they are to be enforced. Thus, the notion that there are "unfettered markets" is a chimera, an idea that is often used by those who are trying to shape markets in a particular way (as if there was a "right way" by which markets should be organized)—often in ways that are in their own interests. Here, we stress the importance of these rules and regulations in shaping the rate and direction of innovation.

In short, if technological progress is endogenous, there is a raft of market failures: Markets are not likely to be perfectly competitive; benefits of research or learning are likely to spill over to others, both today and into the future; firms engaged in research will appropriate only a portion of the societal benefits arising from their research; but attempts to strengthen appropriation will introduce further distortions in the economy.

While the *sources* of market failure are multiple and complex, so are the consequences. Both the level of R & D, the portfolio of R & D research projects, and the direction of research are distorted. Because, as we have noted, production is linked to learning, the level and pattern of production is distorted, relative to the first best. Because labor contracts too affect labor mobility and the extent and manner in which learning occurs, these too are distorted (relative to what they would be in a society that sought to maximize learning). The central thesis of this book is that every aspect of the market economy (and more broadly of our society) needs to be reexamined from the perspective of learning and innovation.

Do Markets Engage in Too Little Research and Learning?

One of the central questions in the economics of innovation and learning is: Do markets on their own engage in too little research and learning? The central insight that knowledge is a public good suggests that the answer is yes. And indeed, the case that markets invest too little in basic research seems compelling—with the obvious implication that there is an important role for government. But what about more applied research? Most of the market failures (e.g., those arising from imperfections of competition and the inability to appropriate all the benefits of R & D) suggest that there is in fact underinvestment in research and learning, certainly relative to what would be the case in a first-best world, and even (as we show more clearly in part 2) relative to a second-best world, in which there are a variety of constraints on the kinds of interventions that government can undertake. But our discussion has also made it clear that matters are more complicated; in some cases, private returns can exceed social returns, in which case there can be excessive research—and especially excessive research in certain areas.

In markets with imperfect competition, one of the objectives of research is rent seeking—obtaining the monopoly rents derived from patents or simply from the first-mover advantage. The fact that investments in innovations are driven by rent seeking shows that rent seeking need not *only* have adverse effects on the economy. It can be channeled toward more constructive purposes.[59] But rent seeking here, as elsewhere, can also result in distortions to the economy, as firms direct research to seize part of the profits of rivals, in me-too inventions.

There are other forces offsetting the tendency for market economies to underinvest in learning and research. Frank Knight (1921) long ago noted the tendency of entrepreneurs to be irrationally overconfident—one might say irrationally exuberant. They systematically believe that the returns on their innovative activities will be greater than they will be, that the probability of failure is smaller. Entrepreneurs *have* to have confidence in themselves and in their relative ability. But if this is so, it means that the level of investment (including investments in R & D and learning), especially in certain "exciting" areas, may be excessive—excessive given the *private* returns, though not necessarily from the perspective of social returns. Indeed, this irrational exuberance serves to partially counterbalance the underinvestment arising from the market failures upon which we have focused in this chapter.

(Some of this seeming irrationality can, in fact, be explained in models with imperfect and asymmetric information, using models analogous to those that explain the winners' curse in auctions. It is those that have obtained the most favorable information that bid the highest;

but then, in formulating their bid, they need to take into account that others have obtained information that is less favorable. So too for the decision to undertake any project, including a research project.)

Intellectual Property Rights

Much of the popular discussion of innovation focuses on the consequences of the imperfect appropriability of the social returns to innovation. Given that this is seen as the central problem, it is natural that attention is focused on government policies at improving appropriability, through strong intellectual property rights.

Our analysis has shown that this focus is incorrect in several respects. First, it focuses on only one of several market failures. We noted that markets where innovation is important are likely to be imperfectly competitive; poorly designed IPR regimes may exacerbate this imperfection.

Secondly, the attempt to correct this problem through strong patent protection can result, as we noted, not only in underutilization of knowledge but in overinvestment, especially in certain types of research. Markets may not only invest too much or too little in research, they may invest too much in some kinds of research (me-too patents in the drug industry or research that may lead to holdup patents) and too little in others (especially in basic research).

In fact, the patent system may itself lead to further distortions in the market—with stronger (and especially poorly designed) intellectual property regimes actually slowing down the overall pace of innovation and increasing the inefficiency of the market's innovation process.

Patterns of Research and Learning

As we have said, it is not just a matter of the *level* of R & D or learning. There may be too little risk taking of some kinds, too much of others; too much attention to correlation under some circumstances, too little in others.

Even more disturbing is that the *direction* of research is distorted. There is clearly too little research aimed at reducing environmental impacts (say, those associated with global warming)—not surprisingly, given the absence of a price associated with carbon emissions—and too little attention paid to the unemployment and distributive consequences of innovation.

Our analysis calls into question Schumpeter's euphoria about the virtues of a market in producing innovation. Schumpeter suggested (though never proved) that competition to be the dominant firm would lead to a high level (the "right" level, perhaps) of innovation. At the same time, he argued that monopoly power would be temporary and checked by potential competition. In chapter 5, we questioned those results: Monopoly power may persist, and the threat of competition, rather than leading to more innovation, may lead to costly entry deterrence. And the entry deterrence can be sufficiently successful that the monopolist can enjoy high profits.

While the high profits do provide a way to finance the up-front sunk costs associated with R & D, particularly important in the context of imperfect capital markets, this is not the best way to finance research, i.e., the way that is most equitable and least distortionary. Also, there are, as always, high costs associated with even temporary monopolies. In some cases, putting aside excess returns to the owners of the monopoly, more of the profits are invested in marketing and advertising than on research, with both marketing, advertising, and research directed more at further enhancing market power (reducing demand elasticities, increasing switching costs, disadvantaging rivals) than at enhancing consumer and societal welfare.

Moreover, Schumpeter, in his support for monopoly, ignored agency effects, contributing to the lethargy that is often associated with monopolies.

Innovation and the Enhancement of Individual and Societal Well-Being

Just as Schumpeter's faith that "Schumpeterian competition" would lead to overall economic efficiency appears misplaced, so too, Schumpeter's optimism that *all* (or most) citizens would benefit from dynamic capitalism appears unwarranted. Twenty-first-century capitalism illustrates that inequality can increase so much that most individuals can be worse off: In the United States, median household income has been falling and, as this book goes to press, is lower (adjusted for inflation) than it was a quarter century ago. And this does not take into account the decreased sense of well-being from increased insecurity and environmental degradation. Those who lost their homes and their life savings as a result of the "innovations" of America's financial system may take little comfort in the notion that *perhaps* their grandchildren will be better off. (The realization that, say, the median income of male full-time workers is lower today than it was forty years ago may also diminish confidence in trickle-down economics.)

Those who glorify the market's innovativeness, of course, pay little attention to the distributive and unemployment effects. They believe that the market (unfettered by government direction) will produce the highest level of innovation and that that will result in the highest level of societal welfare. The presumption is that there may be winners and losers, but society as a whole benefits, which means that the winners could more than compensate the losers. One naïve version of this holds that there are no losers, that somehow the benefits do trickle down to everyone. There is no empirical support for the strong version of trickle-down economics—it should be obvious that, repeatedly, large numbers of individuals have become worse off as a result of innovations that seemingly increased GDP.

A weaker version holds that *eventually* everyone benefits from higher growth, and that would be true if at the same time there was not an increase in inequality. But in recent years, as we have noted, growth in the United States has been associated with marked increases in inequality, so large that significant parts of the population—in some cases a majority—over long periods of time, have seen their standards of living erode, and the increase in inequality may itself be, at least in part, a consequence of innovation, and the way that markets have directed innovative activity. (Matters can be made better or worse by government policy; more recently, some countries seem to have taken the stance that for the country to compete, social programs have to be cut back, so those at the bottom and middle have suffered even more.)

Schumpeter was right that over the two hundred years prior to his writings, innovation had been so strong that almost all benefited. It does not follow that that will necessarily be true over the next hundred years.

It should be obvious that if market prices are distorted, then the market will pay insufficient attention to saving underpriced resources. Because environmental resources are underpriced, innovation is excessively directed at saving labor and insufficiently directed at saving natural resources.

So too, endogenous labor-saving innovation almost surely has played a role in contributing to growing inequality in more advanced industrial countries. Government interventions in R & D that redirect innovation may make it more likely that more will benefit and fewer will lose.

This book is much about how the government either deliberatively or unintentionally directs this process of Schumpeterian creative destruction, by its own research programs (say, on basic research) and the terms on which it makes the results of that research program available; by the structure of the intellectual property laws (e.g., what has to be disclosed, what can be patented, the breadth of the patent and the novelty standard, the nature of the remedy for violation); and through virtually every other aspect of the country's legal and economic framework, including the standards for competition law (when will firms be found in violation of such laws, and what is the enforcement). Each of these laws and policies affects the pattern and direction of innovation, so that all governments implicitly or explicitly have an innovation policy; they simply may not know it. The U.S. legal system led to a policy that almost surely encouraged financial innovations relative to the social optimum and discouraged other kinds of innovation, such as those which would have helped protect the environment (more by what they did not do—ensuring that there were appropriate environmental prices in place—than by what they did do).

The presence of the pervasive market failures associated with learning and innovation that we have detailed in this and the previous chapter raises the question: What would constitute optimal, or at least better, resource allocations to innovation? More broadly, what government interventions would enhance societal well-being? The next chapters of this book provides an analysis of what optimal interventions might look like in the context of some highly stylized models, while part 2 discusses more broadly an array of policy interventions that may enhance societal well-being. Some of these interventions are "fine-tuned," calling for targeted subsidies to one sector or technology. But others are broader in scope and can (and we would argue should) be undertaken even by governments with limited capacities; we show that there is a presumption that developing countries should protect their industrial (including "modern" service and agricultural) sectors.

Markets on their own will not create a learning society, or, even if they do, they will do so more slowly and less extensively than they should. Governments can help infant economies grow. In most of the countries that have been successful in making the transition from less developed to more developed, from a stagnant economy to a dynamic learning economy, governments have done so.

CHAPTER SEVEN

Learning in a Closed Economy

WHILE ONE of the major impetuses for our writing this book was to develop the infant-economy argument for protection, which claims, in part, that some degree of protection can facilitate learning, learning is important in all economies and societies, including “closed” economies. To fully understand the role of learning in an open economy, one has to understand how public policy can be used to enhance learning even in a closed economy.

Chapter 6 showed that, in general, in an innovation and learning economy, the market equilibrium is not (Pareto) efficient. In this and the next chapter, we seek to translate these general principles into more concrete results. We want to show that output in the learning sectors—and in particular, in sectors with large learning spillovers to other sectors—will be too low, and that government intervention, both through subsidies and trade protection, may be desirable. The later chapters then expand on these general principles, exploring the implications for policy in a number of arenas.

1. Basic Competitive Model

In this chapter, we present a simple model to highlight the issues at hand. We assume that there are two types of goods—one industrial or manufacturing and the other agricultural or craft. We assume that the only way that productivity is increased is by "learning by doing," and that such learning only occurs in the industrial sector, but that the benefits of this learning spill over fully to the agricultural sector.

If the government distorts production, by subsidizing production of manufactured goods, societal well-being is decreased *today*. And that has been the traditional argument against such "industrial policies"—policies of government trying to encourage particular industries or particular technologies. But these traditional arguments ignore the future benefits: because of the expansion of the industrial sector, there is more learning, from which both sectors benefit. Well-being in the future is increased.

It is *always* desirable to encourage, to some extent, the industrial sector, so long as the marginal learning benefits are strictly positive. For at the "no intervention" equilibrium, the marginal cost of a small intervention is of second order (i.e., negligible), while the marginal benefit of the cost reduction is strictly positive.

In a closed economy (with no trade), the way in which governments typically encourage the production of one sector is to subsidize it. But to provide a subsidy, one must raise taxes, and there is typically a cost of doing so. Thus, the extent of subsidy that is desirable will depend on the cost of raising funds. If it is possible to levy lump sum taxes, the government should do so, not imposing taxes on the non-learning sectors. But even if the only way to raise revenues to finance subsidies is through distortionary taxation, it is desirable to do so.

These results are more general: by shifting production towards sectors or technologies with high learning and learning spillovers, the economy can achieve higher growth, year after year.

In simple aggregate models (of the kind that have predominated in economics for a half century), there is, of course, no scope for the growth rate to be increased through sectoral rebalancing.

The government's objective is thus to expand sectors with high learning elasticities and high spillovers. Sometimes, that can be done best, by encouraging such sectors *indirectly,* for instance, by subsidizing complements of the good (products such that when their price is lowered, individuals consume more of them—like tea and sugar) and taxing substitutes. The government can also encourage the expansion of such sectors by providing a greater supply of public goods that are complements to such sectors.

Interestingly, even when all goods are symmetric, the market is inefficient. For if individuals worked harder today, there would be more learning, and productivity next period would be higher. But the competitive market will not take this into account. Thus, the government should subsidize work. The magnitude of the distortion will be greater the lower the discount rate (the more valuable future consumption), the higher the learning elasticity, and the easier it is to induce workers to work more. This labor supply effect is, of course, relevant also in non-symmetric models: sectors which are complements to work (substitutes for leisure) should be encouraged through, for example, lower taxes.

There is a more general point: one can only assess the distortions arising from a market failure within a general equilibrium model. If the production of some sector is smaller than it otherwise would be, that of some other sector is larger. In the presence of learning, learning sectors have a lower level of output, non-learning sectors have a higher level of output, and, overall, output is likely to be too small. The competitive equilibrium will be associated with lower rates of learning than is optimal.

Optimal Interventions

We noted earlier that a small intervention—a slight increase in the production of manufactured goods relative to the competitive equilibrium—has a negligible effect on consumer well-being the first period. But the outward shift in future production possibilities curves has a first order effect on consumer welfare. That was why it was optimal to have *some* intervention. But as we increase the size of the distortion, the marginal costs increase, and the marginal benefits (typically) decrease. This means that there is an optimal intervention—an optimal subsidy.

The lower the discount rate—the more we value future growth, and the higher the learning benefits (the greater spillovers to other sectors)—the more we should want to distort production. In a model with *full* learning spillovers, the distortion occurs only in the industrial sector, while the benefits are economy-wide. Hence, the smaller the share of the manufacturing sector, the larger the desired distortion. In Appendix A, we provide a simple formula for the optimal subsidy.

Multiple Equilibria and the Possibility of a Low Level Equilibrium Trap

The benefits of learning are related to the future size of the economy. The larger the economy, the larger the cost savings from any reduction in the costs of production. The larger output today, the lower the costs of production next period and the larger the size of the economy next period. It is possible that there may be multiple equilibria. With higher expected output next period, it pays to produce at a high level today; and given that high level of production, costs next period are low and production is high. The expectations are justified. But it is also possible that there is a low level equilibrium trap where production is low today in the expectation that output will be low next period; and because production is low, learning is low, and output will in fact be low next period.

When there are multiple equilibria of this sort, government intervention can help move the economy from the low growth equilibrium to the high growth equilibrium.

2. Monopoly

Sectors with learning (or where innovation is important) exhibit increasing returns and are natural monopolies. At the very least, they are likely to be characterized by limited competition. Earlier chapters showed that monopolies may have an advantage over competitive markets in that they internalize the learning externality. They are aware that if they produce more, there will be more learning, and future costs will be lowered. And because they are aware of these future gains, they expand production of the learning good *toward* the level that optimal government intervention would have called for. This "internalization effect" countervails their normal exercise of monopoly power, which entails *restricting* production.

But while in economies with monopoly/monopolistic competition, there is *some* internalization of learning benefits, the firm not only still does not internalize cross-sector learning effects, it also does not take into account demand effects; that is, the actions of firm j (controlling product j) affect the demand for product i, and therefore affect the equilibrium level of innovation in that sector (even in the absence of learning spillovers). Moreover, there are macro labor-supply effects (described earlier) that firms fail to take into account. Government subsidies can help correct these distortions.

What makes the analysis still more complicated, though, is that even in the absence of learning effects, the general equilibrium with monopolistic competition (either with a fixed or variable number of firms) will not, in general, be Pareto efficient. Dixit and Stiglitz (1977) showed that with free entry, with all firms facing the same fixed costs of entry, symmetry, and constant elasticity of demand curves (and no learning), the market equilibrium was efficient. However, when these idealized conditions were dropped, it was not.[1]

Still, we can say something about the nature of monopoly distortions as well as desirable forms of government intervention. In the discussion that follows, it is important to differentiate between the monopoly/monopolistic competition equilibrium and the competitive equilibrium, with and without optimal subsidies. Innovation will be lower in the monopoly equilibrium than in the competitive equilibrium with optimal subsidies. The more subtle question is whether the unsubsidized competitive market is more innovative than the unsubsidized monopoly.

If the elasticity of demand is relatively low, it is more likely that the demand restricting effect will dominate. If the discount rate is low and the learning elasticity is high, it is more likely that the "internalization" effect will dominate. If the internalization effect is large enough, then the benefits of learning dominate. In other words, future consumption is increased so much from the lower costs of production that the gains in future utility outweigh the losses from the exercise of monopoly distortion.[2] It is thus *possible* that output of the learning sector with monopoly is higher (initially) than in the competitive equilibrium. When this is not the case, the economy suffers from both the short-run distortion and the slower, long-run growth. When this is the case, the economy initially benefits from being closer to the social optimum. But unless productivity is expanded enough as a result of this greater learning, output and utility next period would be lower, were it not for the learning benefits that persisted. But if the monopolist again internalizes these learning benefits, output next period would be higher. Indeed, because input and growth are higher than in the competitive equilibrium, eventually even the level of consumption—restricted as it is by the exercise of monopoly power—will be higher, even in the absence of the internalization of learning benefits.

We can ascertain some conditions under which it is more likely that learning with monopoly will be higher than with competition. *The greater the elasticity of labor supply, the more likely that the monopoly innovation level will be greater than the competitive level. On the other hand, with a sufficiently low elasticity of demand, low level of learning and learning elasticity, and high discount rate, the level of innovation with monopoly is lower than that with competition*.

A higher level of learning in monopoly may still be associated with a lower level of utility. Even though the production possibilities curve has shifted out with learning more than would have been the case under competition, and consumption in earlier periods is higher, the distortions associated with monopoly in later periods when learning may be less important imply that the level of utility attained by consumers can be lower. Thus, we have three different possibilities:

1) With low learning, the traditional result—lower consumer welfare both periods and lower innovation—holds. 2) With high learning, consumer welfare may be higher both periods, so consumers are clearly better off. 3) In the intermediate case, innovation may be higher, but if consumers are better off the first period, they are not enough better off to compensate for the loss of welfare the second period as a result of the exercise of monopoly power—even though the production possibilities curve has shifted out.

In the symmetric case of monopolistic competition, where there is no distortion in the relative consumption of different goods, there is still a distortion in labor supply (as was the case in the model of competition): the exercise of monopoly power lowers real wages, and this (normally) leads to a lower labor supply, less output, and less learning. This is always the case with myopic monopolies (which do not take into account the benefits of future learning). *With myopic monopolies, in the symmetric case, monopoly is unambiguously worse than competition; even though neither takes into account the benefits of learning, growth is higher with competition than with monopoly.*

A non-myopic monopoly internalizes the learning benefit, as we have noted, and this makes the results more clouded. As noted earlier, monopoly can be better if the learning elasticity is high, the discount factor is low, and the demand elasticity is high.

If the government can undo the monopoly power and intervene with optimal subsidies, it should obviously do so. But if it cannot—if it has to live with the monopoly—then it should still subsidize the monopoly (setting aside any distributive concerns). Because a monopolist will charge a price which is a mark-up over marginal costs, a per-unit subsidy has a larger effect in lowering prices, increasing output, and thus advancing learning, than has a comparable per-unit subsidy in a competitive market. There is some ambiguity about the size of the marginal benefits of a given per unit subsidy; since production next period is lower than it would have been in the competitive-equilibrium-with-optimum subsidies because of the monopolist's restrictions on production (assuming that this adverse effect is not fully offset by the internalization of the learning in later periods), the marginal benefit from any increase in learning is lower; but a subsidy has the additional benefit of helping correct the monopoly distortion, which leads to an underproduction of the manufactured goods.

3. Concluding Comments

This chapter has pursued three of the key insights of the chapters of part 1: (a) the industrial sector is likely to be more innovative with important spillovers to other sectors than, say, agriculture; (b) unless there are perfect spillovers, markets will not be competitive; and (c) whether competitive or not, markets will not be efficient. We have shown (a) under competition, the industrial sector will be too small, and subsidies are desirable—even if the funds to pay the subsidies are raised through distortionary taxation; (b) under monopolistic competition, the industrial sectors may be larger or smaller than in the competitive equilibrium, depending on the relative importance of the benefit from learning internalization; (c) under monopolistic competition, if there are cross-sector learning spillovers (distribution effects aside) it will be desirable to subsidize the industrial sectors—to encourage them to expand; (d) there are inefficiencies not only in the composition of output but also in labor supply; it will normally be lower than is optimal because of the learning benefits from greater production.

With precise knowledge of the demand, production, and learning functions it is possible to design "optimal interventions," balancing the short-run distortions with the long-run benefits. Even without such detailed knowledge, if we can identify a set of sectors, like the industrial sector, with higher learning elasticities and higher learning externalities, we should encourage such sectors. Later chapters will show how it is possible to do so through broad-based measures like changes in the exchange rate.

Schumpeter argued for the advantages of monopoly. In the context of our learning model, the central advantage is that they internalize the learning externality. But as we noted, there is a large cost: a tendency to underproduction. In the case of investments in R & D, that implies less of an incentive for research. In the case of the learning models on which we focus in this book, it means less learning. We have delineated conditions in which the benefits of the internalization of the within-industry learning externality outweigh the costs of the monopolistic distortion of the market.

In any case, Schumpeter's enthusiasm for monopoly is not justified, for one can do better by government intervention—imposing a subsidy on the learning sector, paid for either by tax on the other sector or, if feasible, by a lump sum tax.[3]

APPENDIX A

Diagrammatic Exposition

The basic ideas outlined in the text can be illustrated by some simple diagrams. Assume that each of the two goods (manufacturing and agricultural) are produced just by labor, at fixed input per unit output each period. We let C_A be the consumption of agricultural goods, C_M of manufactured goods. With a fixed labor supply, the production possibilities curve is a straight line as depicted in figure 7.1.

We assume that learning only occurs in the manufacturing sector, that the more production, the more learning, and that there are full spillovers to the other sector. Full spillovers means that the production possibilities schedule has the same slope next period as it does this period.

In competitive equilibrium, where no firm pays any attention to the learning benefits, the competitive equilibrium is the tangency between the indifference curve and the production possibilities schedule, and is denoted by $\{C_A{}^{t*}, C_M{}^{t*}\}$. There will be some learning, moving out the production possibilities curve, so the next period (denoted by $t+1$)

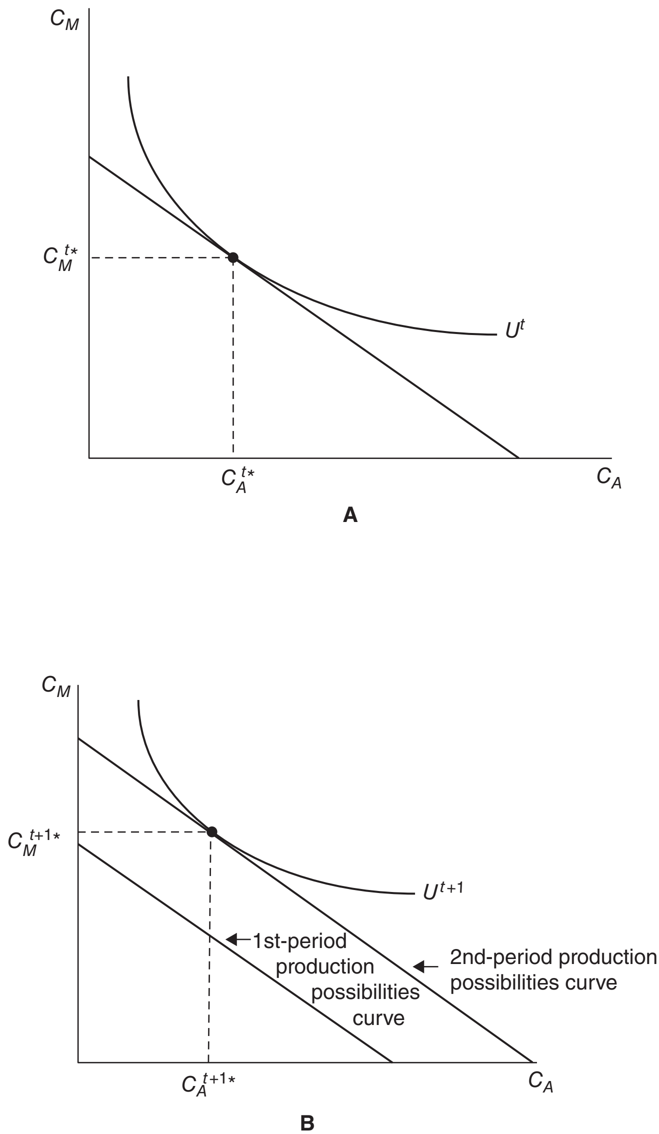

Figure 7.1 The competitive equilibrium.

A. First-period equilibrium occurs at the tangency between the production possibilities schedule and the indifference curve, at C_A^{t*}, C_M^{t*}. B. Second-period equilibrium looks much like the first, except that the production possibilities schedule has shifted out. The equilibrium occurs at C_A^{t+1*}, C_M^{t+1*}.

there is more consumption of both goods, as depicted in the lower panel of figure 7.1.

Figure 7.2. shows what happens when the government subsidizes manufactured goods at time t. It distorts production, lowering utility at time t, but there is more learning, so that the production possibilities curve shifts out—increasing welfare at time $t + 1$. For small subsidies, the first period distortion is small (of second order) relative to the learning benefit: it is always optimal to have some subsidy.

Figure 7.3 shows the effect of monopoly in the manufacturing sector. With no learning, the production of the manufactured good is lower than in the competitive equilibrium (more resources are deployed in agriculture). This effect is offset by the internalization of learning. In the case depicted, the internalization effect outweighs the monopolization effect, and the production of manufactured goods at time t is greater than in the competitive equilibrium. This means the production possibilities schedule at time $t + 1$ is shifted out (relative to what it would have been in the competitive equilibrium). Because of the monopolization effect, at $t + 1$ utility is not as high as it could have been *given the better production possibilities schedule*. Still, utility is higher than it would have been in the competitive equilibrium. It is, of course, still ambiguous whether overall utility has increased. In the case depicted, where

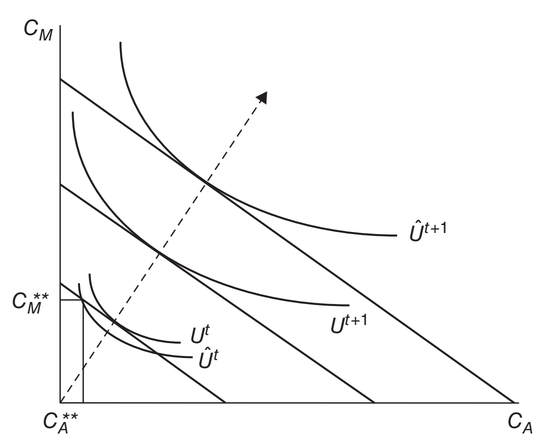

Figure 7.2 An increase in C_M^t above C_M^{t*} lowers utility in period t slightly, but leads to a large increase in second-period utility. $\hat{U}$ represents that levels of utility each period as a result of the distortion from the competitive equilibrium.

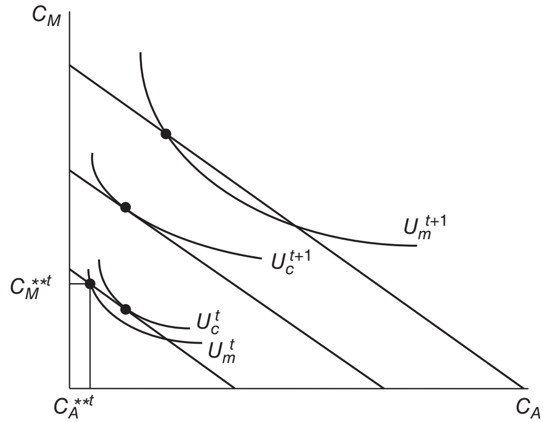

Figure 7.3 A monopoly may increase its production (relative to the competitive equilibrium) in the first period because it takes into account learning benefits. This means that the production possibilities curve in the second period is further out, so much so that utility in the second period is increased enough to offset the losses in utility during the first period from overproduction of industrial goods, relative to the static equilibrium. In this case, monopoly is better than competition. U_c represents the utility in the competitive equilibrium.

the learning benefits are large, overall utility would be increased, unless future consumption is discounted very heavily.

APPENDIX B

Optimal Interventions

One can derive simple formulae describing the optimal subsidy for manufactured goods. The most general result is parallel to that of Ramsey concerning optimal taxation: the percentage reduction of consumption (along the compensated demand curve) for each good should be proportional to the *total* marginal learning benefit from encouraging the consumptions (equals the production in a closed economy) of sector *i*, taking into account effects on other sectors, both through induced learning in other sectors (as a result of cross-elasticities of demand) and as a result of learning spillovers. While the general expressions are complicated, in the case of separable demand functions (where the demand for each good depends only on its own price), and with two goods, we obtain

$$t_M = -\rho h \nu / s_M$$

where

$$\nu = L^{t+1}/L^t$$

the ratio of labor input in the two periods; where h is the elasticity of the learning curve (the percentage reduction in next period's cost as a result of a percentage increase in output this period); where ρ represents the amount by which future consumption is discounted;[4] and where s_M = fraction of labor allocated to manufactured goods in the first period.

We have obtained a remarkably simple formula for the optimal industrial sector subsidy.[5] The industrial sector subsidy should be greater the higher the value of future income (ρ), the higher the learning elasticity (h), and the lower the share of industrial goods in production. The benefits of learning relative to the distortionary costs will be greater the greater the relative scale of the economy in the future (i.e., the greater ν).

Perhaps the most surprising result is that the optimal subsidy does *not* depend on demand elasticities (quite unlike the standard result in optimal tax theory). The reason is that both the costs (the distortion) and the benefit (consumer surplus associated with lowering the costs of production) are related to the (compensated) demand elasticities, in a fully offsetting way.

CHAPTER EIGHT

The Infant-Economy Argument for Protection: Trade Policy in a Learning Environment

CONVENTIONAL ECONOMIC wisdom is that free trade enhances economic efficiency and this promotes growth. Indeed, there are few propositions in economics about which there is greater consensus among conventional economists than those that assert the benefits of free trade.[1, 2]

However, in the "learning" context upon which we focus here, spillovers (both technological and institutional) within countries but across industries may be fundamental to the process of growth. As we have already noted, there will be trade-offs between static (in)efficiencies and dynamic benefits. In the last several chapters, we saw how production in earlier periods (especially in sectors with high learning elasticities and learning spillovers) was increased (beyond the level associated with static efficiency) in order to increase output in later periods. In an open economy, the essential trade-off is between the static efficiencies associated with comparative advantage and the dynamic benefits associated with the faster learning that might come from alternative resource allocations, including those associated with beneficial local spillovers.[3]

It is desirable for governments to intervene in the market, to encourage, possibly through trade protection, sectors in which there is more learning and more learning spillovers. We call this the infant-*economy* argument for protection: Learning benefits are likely to be especially high for countries for which there is a large knowledge gap between themselves and the more advanced countries.

Before presenting the argument in detail, we discuss a long-standing argument for protection that might seem similar, but in fact is not: the infant-*industry* argument for protection.

1. The Infant-Industry Argument for Protection

The infant-industry argument for protection held that developing countries should protect their "infants" so they could become more productive (learn-by-doing) and thus become competitive with the more advanced countries. Without such protection, the developing economies would be relegated to producing traditional goods marked by slow growth in productivity.[4]

As we argued in chapter 4, the industrial sector is subject to faster learning than the agricultural sector,[5] so it was natural that developing countries would want to move into that sector. But industry was not their current comparative advantage; without some government intervention, they could not enter the industrial sector and therefore could not learn. Unfettered markets, it was feared, would keep a country from entering more dynamic sectors, especially if the learning is external to the firm (for then no firm would have an incentive to make the investments required to "catch up").

With protection, firms could enter these more dynamic sectors. As a result, for instance, of learning by doing, marginal costs would decline, and firms would eventually become competitive. Without protection, they could not survive to enjoy the benefits that come from that learning.

There is, in fact, considerable evidence for the validity of these infant-industry arguments—protection did play a critical role in the development of many European countries (see, in particular, Chang 2002). But critics point out that in many cases, the infants never seem to grow up, imposing high costs on society. We shall address these concerns later in this book—we believe that they are far less relevant for the infant-economy argument for protection than for the infant-industry argument. Indeed, we will explain why government subsidies may be desirable even if the subsidized sector *never* becomes fully competitive.

Moreover, these critics contend that if it were the case that these firms would eventually be competitive, they should be able to borrow today (see, e.g., Baldwin 1969). It is easy to show that in a long-run dynamic model (i.e., where firms maximize the present discounted value of their profits, taking into account future learning), the relevant marginal cost of production is *lowered* by the fact that producing today lowers the marginal cost that they will face in the future.[6] Thus, it would be optimal for them to operate at a loss today. Opponents of protection say that if these firms are really competitive, then they *should* be able to borrow to finance their learning—there is no need for government intervention.

Capital Market Imperfections, Imperfect Information, and the Infant-Industry Argument for Protection

However, especially in developing countries, firms cannot borrow, especially on the basis of future profits (not collateral). The theory of imperfect and asymmetric information has explained why that is so.[7]

The response to this (according to the critics of the infant-industry argument) is that government should then step in to correct *this* market imperfection, rather than creating a new political economy problem. If that can't be done in general, the government should (in this view) simply lend the money that the firm would have been able to borrow had capital markets worked better, at a commercial interest rate.

To some extent, this is in fact what the successful East Asian countries did (see Stiglitz 1996; Stiglitz and Uy 1996; World Bank 1993). They made capital more available to firms that they believed showed more learning potential. But they went beyond just correcting the failure in financial markets.

There are two reasons that simply correcting those failures will not suffice. The first, which is the subject of most of this book, is that markets and economies in which innovation (learning) is important are rife with market failures. Correcting one market failure—access to finance—leaves in place a rash of others, some of which are far more important. In the case of the East Asian countries, they were not just correcting financial market failures; they used access to scarce finance as an instrument of industrial policy, to encourage sectors with large technological spillovers.

The second response to the suggestion that firm subsidies or loans be used *instead of* (rather than in addition to) protectionism is that this approach does not fully take into account the information imperfections that give rise to capital market imperfections. The government may not be in any better position than the capital market is to judge which *firms* are likely to repay their loans. (As we explain, assessing which industries or sectors have more learning potential and more learning spillovers requires quite different information.)[8]

In a sense, there is some, but not complete, similarity between the use of the patent system to finance research and the use of protection to finance new industries. Government funding of research can be much more efficient, avoiding some of the static and dynamic inefficiencies associated with the patent system (see chapter 12). But the government then has to decide *which* researchers should be funded, just as capital markets have to decide which entrepreneurs to finance.[9]

Both the patent system and protection, though, allow for self-selection. With the patent system, firms take their own chances; those that are lucky or smart will get a return. If firms misjudge, they bear the consequences. So too for protection, or, as we explain later, with other broad-based interventions, such as exchange rate adjustments.[10]

Why Not Entering the More Dynamic Sectors May Not Be a Disadvantage

This market failure we have just described, by itself, does not provide a fully coherent argument for trade and industrial policy. Underlying the infant-industry argument for protection is the belief that countries are disadvantaged if they cannot enter sectors with high productivity growth. This is, however, not necessarily the case.

The reason is that all countries will benefit from the learning in the more dynamic economies as a result of lower prices, so long as markets remain competitive. Moving into a more dynamic sector does not guarantee a country greater (innovation) rents. And those countries not in the sector benefit from the learning going on elsewhere, *so long as global markets are competitive*.

With competition, prices in the dynamic sector fall in proportion to productivity, so that while revenue per unit produced falls in proportion to productivity, revenue per hour remains the same. The economies with the dynamic sector benefit from learning with the reduction in costs. The countries specialized in the non-dynamic sectors should not envy the other countries that have specialized in the seemingly more dynamic sector. They benefit from the lower prices they have to pay for the goods produced by that sector.

This analysis, however, assumes away two of the critical market failures in markets with learning and innovation that we highlighted in chapters 5 and 6: learning spillovers and imperfections of competition. Each of them alone provides a convincing rationale for trade and industrial policies, and the two of them interact to provide an even more compelling case for government intervention (and even more so in conjunction with the other market failures delineated earlier in the book).

2. From the Infant-Industry Argument to the Infant-Economy Argument

This book stresses the importance of creating a *learning economy*. Our focus is thus not on particular sectors but on the broader economic system. In this chapter, we argue that protectionism can be an important instrument for helping infant economies grow by creating a learning society.

We provide the real answer to the critics of industrial and trade policy—a fully articulated rationale for protection based on market failures derived from information asymmetries and endogenous learning and innovation *with learning externalities*.

While earlier chapters (and earlier discussions within this chapter) highlighted that there were multiple market failures *inherently* associated with innovation (including imperfections in risk markets and in competition), this chapter will focus more narrowly on learning externalities.

In this perspective, industrial and trade policy is not focused on picking winners, though, to be sure, governments do not want to pick losers. Nor is it predicated on the belief that government can do a better job than the private sector of picking winners. *It is based on the notion that learning involves spillovers (externalities) that will be imperfectly internalized in a market economy*. Industrial and trade policies are concerned with identifying sectors or industries (firms, areas of innovation) which would generate large externalities or where the returns that could be appropriated by the innovating (learning) firm are a fraction of societal benefits. Governments in many countries have, in fact, done a credible job in making these selections, and our societies have benefited greatly as a result. Chapter 9 discusses these historical experiences.

The infant-economy argument for protection that we advance here does not even require identifying particular sectors with large spillovers or large capital-market imperfections. It simply argues that *on average* spillovers may be larger within some broadly defined sectors of the economy—sufficiently larger to warrant distortions in the conventional static allocation of resources.

In order to show more fully the role of trade policy, we continue the analysis of the simple two-sector model with an industrial (modern) and a traditional ("craft" or "agricultural") sector introduced in chapter 7, extending it to an open economy. Recall from that discussion that there are four key features to the model: (a) there are spillovers from the industrial sector to the crafts sector, for which firms in the industrial sector are not compensated; (b) such spillovers are geographically based, that is, only productivity increases in the industrial sector in the developing country affect productivity increases in the traditional sector;[11] (c) the industrial sector is the sector in which innovations are concentrated; and (d) among the important determinants of the pace of innovation in the industrial sector in the developing country (or of its impact on the traditional sector) is the size of that industrial sector.[12]

Earlier critiques of trade policies encouraging the development of the industrial sector in developing countries ignored the spillovers that are at the heart of the analysis here. They argued, first, that protection is costly: Korea could have more industrial goods and more agricultural goods by taking advantage of its comparative advantage.

Second, critics of protectionism contended further that Korea's comparative advantage wouldn't change as a result of protectionism. Korea would always have a comparative advantage in growing rice. Therefore, it was foolish for it to restrict imports of industrial goods, even if by doing so increased productivity in the industrial goods sector. It could never catch up, so the protection would have to be permanent. Year after year, the country would have been better off if it simply specialized in its own comparative advantage, growing rice.[13]

There are two errors in this line of reasoning. The first is that countries can and do catch up, at least in certain areas, and Korea provides a telling example. If catching up is possible, then *dynamic* comparative advantage differs from static comparative advantage.[14] Most importantly, *if dynamic comparative advantage is affected by today's resource allocation, then it is desirable to intervene in the market, to move away from static comparative advantage*.

But the second flaw is even more profound. The standard analysis ignores learning spillovers and the fact that there may be more learning, and more learning spillovers, associated with some sectors than others. *If there are advantages to industrialization (e.g., associated with learning and learning spillovers), as our earlier analysis suggests, then again it is desirable to intervene in the market, to move away from static comparative advantage*. Even if the infant never grows up—even if year after year there has to be a subsidy, even if there is no change in comparative advantage—the dynamic benefits of protection, and the faster rate of growth that results, may (and under our assumptions will) exceed the static costs. Intervention may be desirable, because it enhances an economy's learning and its ability to learn. That is, even if Korea's comparative advantage remained in agriculture, industrial protection might be desirable, because by doing so, one might also have a more dynamic agricultural (traditional) sector, as a result of the learning spillovers from the industrial sector. Trade restrictions enhance the size of the industrial sector, the benefits spill over to the rural sector, and national income grows at a possibly far faster pace. Our concern is thus maximizing the learning not just of a particular firm or a particular industry but of the entire economy—creating a learning society.

Whether the short-run costs are worth the long-run benefits depends on the pace of learning and the rate of discount. Of course, countries don't have to choose between the extremes of autarky and free trade. The discussion that follows will identify the nature of the *optimal* trade intervention.

3. A Simplified Model

We consider a highly simplified world consisting of two economies—one developed (D) and the other less developed (L). (The limitation to just two economies is inessential; our model could equally well consist of multiple [identical] versions of each of the two types of economy.)

These economies produce two types of goods—one industrial (M) and the other agricultural/craft (A). (Again, we could easily extend the model to include a multiplicity of goods in each category.) Both are produced using only labor as an input, with technologies that at any point in time embody constant-returns-to-scale.

We assume that the developed economy enjoys absolute advantages in the production of both goods, but that the less-developed economy enjoys a comparative advantage in agricultural/craft production.

We further assume that the developed economy is very large relative to the less-developed economy; in particular, that it is capable of supporting the entire global demand for industrial output and at the same time producing significant amounts of agricultural/craft output.

Free-Trade Competitive Equilibrium

Because of its comparative disadvantage in producing industrial goods, industrial production in the less-developed economy is not economically viable. It specializes in agriculture. The composition of consumption in the less-developed economy is then determined by the relative cost of production in the developed country. The composition of output in the industrial economy is determined by the global demand (its own demand plus the imports of the less-developed economy) for industrial goods. Finally, note that, in this simple static equilibrium, all the gains from trade accrue to the less-developed economy.

Dynamic Development

We now introduce technological progress into this static equilibrium. Formally, we will assume, first, that productivity improvement affects the industrial and agricultural/craft sectors equally (the case of perfect spillovers).

Productivity increases spillover fully from one sector within the economy to the other. This has one important simplifying implication: productivity growth does not affect the price of industrial goods relative to agricultural/craft goods. (Our results require only that there be some spillovers from the industrial to the traditional sector within a country.) This means that there is no change over time in comparative advantage.

To repeat the arguments made earlier in the book, productivity growth results from (1) research and development efforts, that, while originally devoted to one sector, have benefits that inevitably spillover to other sectors; (2) human capital improvements, which again, while they arise in one sector, inevitably migrate with labor to other sectors of the economy; (3) the accumulated knowledge and attention of managers and engineers, which, although developed in one sector, also naturally migrate to other sectors; and (4) institutional developments, which, while they arise in response to the needs of one sector, have benefits to others.

Next, we assume, as earlier, that the industrial sector is more conducive to learning than others. Over time, in the absence of protection, the less-developed economy, specializing in agriculture, falls further and further behind its developed counterpart.

The Role of Trade Policy

Consider now the consequences of a ban by the less-developed country on industrial imports (or equivalently, the imposition of prohibitively high tariffs). The result would be an immediate welfare loss as it substituted high-cost domestic industrial production for lower-cost imports from the developed economy. However, in the new autarkic equilibrium, industrial output in the less-developed economy would no longer be zero, and productivity growth would now occur. Just as in the case of the developed economy, a high-tariff, less-developed economy would produce a mix of outputs dependent on its own demands for industrial and agricultural/craft products at the relative cost of production of the two goods.

Eventually the benefits of this dynamic improvement in productivity will outweigh the short-term inefficiencies associated with high-cost local industrial production. The country will be better off. Whether the present discounted value of welfare is higher depends on the growth rate with protection and the discount rate. Thus, in this context, trade barriers may enhance rather than impair economic welfare. With a low enough discount rate, autarky is always better.

In this model with full spillovers, the developing country *always* has a comparative disadvantage in the learning good; hence, if it is to continue learning, protection must continue. The infant never grows up, in the sense that agriculture remains the country's comparative advantage—and yet it is desirable to continue to provide subsidies to the learning sector.

4. Optimal Trade Interventions

Our analysis can be used to derive an optimal tariff balancing the long-term benefits of fostering industrial growth against the short-term costs of inefficient acquisition of industrial products.

Quotas

Consider a quota on the import of manufactured goods. The country on which the quota is imposed may have a demand in excess of the quota. The quota has a cost: the country is less efficient in producing manufactured goods than the developed country. The quota also has a benefit: the learning improves productivity in both sectors. The government can balance the benefits of further tightening the quota with the costs. A quota just slightly less than what the country would have imported under free trade has a relatively small cost, but the costs increase with tightening of the quota. Just as we saw in chapter 7 that some subsidy is desirable, so too here; *a small quota is always welfare improving*. The loss of welfare from the distortions associated with trade intervention (e.g., the reduction in the consumption of nontraded goods) has, at the market equilibrium (which is optimal, ignoring learning benefits), second-order effects, while the learning benefits from increased production of manufactured goods have first-order effects.

While the costs of intervention increase with the size of the intervention, so do the benefits. But what matter are the *marginal* benefit and the *marginal* cost. The optimal intervention entails the marginal cost equaling the marginal benefit. Growth increases with the size of the trade intervention, but there are diminishing returns. At the same time, there are increased marginal costs associated with larger and larger distortions. The growth benefits at the margin will be higher the lower the discount rate is and the more productivity responds to industrial output. Thus, the optimal quota will be smaller the less future income is discounted and the greater learning benefits are. In the limit, if there are no learning benefits, we obtain the standard result that there should be no trade restriction.

Tax Interventions

We can also subsidize domestic manufactured goods production and tax imports of manufactured goods and domestic production of agricultural goods (similar to the optimal taxes and subsidies discussed in the previous chapter).

It is debatable whether it is possible to provide such subsidies and taxes (or impose quotas, as in the previous subsection) without contravening international trade rules. But sometimes developing countries can impose domestic consumption taxes with much of the same effect. Many countries, for instance, import luxury biscuits at the same time there is domestic production of lower quality biscuits. A tax on the consumption of high quality biscuits will encourage the production of domestic biscuits. (But if the country believes there is more learning associated with the production of high quality biscuits than low quality biscuits, this strategy will be counterproductive.)

Exchange Rates

Restrictions on the use of these other instruments provide part of the explanation for why many successful emerging countries have increasingly turned toward managing the exchange rate as a key industrial policy. Consider China. A lower exchange rate has made China's manufacturing sector more competitive. It has also disadvantaged China's agricultural sector; the prices of competing imports are lower—or would be if China did nothing about that sector. The global market for agricultural goods is vastly distorted by Western subsidies. International agreements have allowed subsidies for agricultural goods to continue, even though those for manufacturing are circumscribed. But China's exchange rate policy can take advantage of that; it too can provide subsidies to its farmers without contravening WTO strictures. Of course, such subsidies are costly and take away money that could be better spent on development.

But the benefits from the more competitive exchange rate outweigh the adverse budgetary effects from agricultural subsidies. (There are also potential adverse effects on the non-traded sector.)

The lower exchange rate has a curious effect: it results in China's exports exceeding its imports—a trade surplus. This is the opposite of the conventional wisdom in development. Poor countries should be borrowing from the rich, so poor countries should have a trade deficit that is offset by a capital inflow. China has been lending to the United States and other western countries at a *negative* real interest rate. This might seem curious until one takes the learning benefits into account; the benefits from the increase in China's productivity, resulting from becoming a manufacturing behemoth, more than offset these potential costs.

In fact, in a simple infinite horizon model in which each period looks like the previous one (except that productivity may have increased), it may pay a country to maintain a trade surplus *forever*. Its reserves would accumulate over time. It might seem that the cost of holding these reserves—not using them to invest in, say, infrastructure or technology—would eventually dominate. But as the country contemplates lowering its reserves, it realizes that doing so would lead to lower levels of production of manufactured goods, lower levels of learning, and lower standards of living in the future. It never pays to do so.

In chapter 11, we will further discuss exchange rate policies—both as an instrument of macroeconomic stability and as an industrial policy, assessing both the costs of such measures as well as the benefits.

5. Non-Steady-State Analysis

In the more general case, as the country closes the knowledge gap between itself and the advanced industrial countries, the marginal benefit of learning may decrease, and the marginal cost of the static distortion and the opportunity cost of not using the ever-increasing surplus may increase. The country might then want, at first, to bring its trade surplus down to zero—to stop growing reserves—and then to consume its accumulated surplus.

Other factors would also, of course, affect the country's desired level of surplus. A country with an aging population might want to put aside savings and then, as the aging population enters into retirement, reduce that surplus. Such demographic transitions are not analyzed well in steady-state models.

Out-of-steady-state analysis requires taking into account *changes* in the exchange rate over time—how current policies affect exchange rates and how changing exchange rates affect intertemporal trade-offs. Because the exchange rate in later periods, when the country is more advanced and when it is running a trade deficit, is likely to be higher than in earlier periods, the country experiences a capital loss on the cumulated reserves. This lowers the dynamic benefit of the surplus, implying that it is desirable to have a lower level of learning (and surplus) than would otherwise be the case.

So too, the analysis needs to be modified if learning is related to the level of investment.[15] Then the government will want to increase the capital intensity of production. If in early stages of development, capital goods are imported, it may wish to have a high exchange rate in early years, followed by a lower exchange rate in later years, as the investments bear fruit and the country starts to export. This will result in high returns to investing in early years, and consequently high levels of learning and growth.[16]

We saw in chapter 6 that the optimal "innovation" policy for a country on the knowledge frontier is likely to be different from one for a country that is trying to catch up. By the same token, as the country succeeds in catching up and the distance between itself and the advanced countries narrows, policies may change. It may be optimal, however, for it to remain a follower—never to fully catch up. The upfront costs of catching up outweigh the benefits. Knowledge eventually filters down. Of course, even then, the "follower" will need to respond to what the leader does. If the leader engages in a more active industrial policy which enhances its growth rate, then the follower(s) should likely undertake a more active "learning" policy, to ensure that it does not fall further behind and to take further advantage of the new learning opportunities that are filtering down from the advanced country.

6. Imperfections of Competition

Chapter 5 emphasized that markets in which learning (innovation) is important are likely to be imperfectly competitive, and prices may be (significantly) above average and above marginal costs.[17] Imperfect competition provides, in general, grounds for government intervention.

However, governments in exporting countries may not have incentives to enforce competition among their exporters. The government in, say, the developed country in which the incumbent monopoly is located has to assess the loss of profits to the monopolist and the loss in tax revenues that might accrue to the government from taxing those profits[18] against the gains to *its* consumers. It doesn't weigh the benefits to consumers in other countries; indeed, since it may garner for itself a share of the profits earned by the monopolist through taxes, it even benefits from the monopoly. Even if the government were not beset by "political economy" problems—undue influence from the incumbent monopolist—it might not be in its best interests to encourage more competition.

Because the country of the incumbent already enjoys the learning spillovers from the presence of the industry, the benefits it receives from helping a competitor get established even within its territory are also smaller than the benefits that would accrue to the establishment of a competitor in another country. If there is to be a subsidy to establish a competitor, it would rather free ride on the efforts of others.

We have explained why a developing country may have greater incentives for having an industrial policy (through subsidies or protection) than does the developed country. While such policies might not make sense for the developed country, they do for the developing. So too, the incentive to have strong competition policies may differ. A developed country might even permit a merger of two of its own companies that would reduce competition; it might gain from the greater monopoly rents (which will be shared with the government through the tax system), even if as a result, consumer prices increase and learning decreases. However, the developing country shares in none of the benefits but bears the costs. Hence, it may be desirable for it to oppose such mergers, demanding at the very least divestiture, perhaps to a locally owned company, that would enable it to begin sharing in some of the learning benefits.

(These issues arise not only between developing and developed countries but also among developed countries. Thus, it is no surprise that the EU often takes a more stringent attitude toward anticompetitive behavior on the part of American companies than does the U.S. government.)

Like many of the policies considered in this book, there are important intertemporal trade-offs. Trying to create viable domestic competitors to foreign monopolists and oligopolists has a cost. With direct subsidies, the costs are borne by taxpayers. With protectionism, the costs are borne by consumers. But if successful, there will be a more competitive marketplace in the long run. In the case of protectionism, part of the trade-off may be less competition in the short run, but more in the long run.[19]

In short, the social benefits the developing country will gain from intervening in the market and putting one or more firms in a position to be effective competitors may well be greater than the costs. Through protectionism, they may lose in the short run (although in the case of subsidies, consumers may gain even in the short run), but in the long run, the country gains from the learning spillovers, and the country may even be able to seize a fraction of the producer rents that accrue to foreign producers as a result of imperfections of competition.[20]

Establishing a more competitive marketplace is, of course, a global public good from which consumers everywhere benefit. However, because it is a global public good, there are likely to be insufficient "investments" in creating a more competitive global marketplace.[21]

7. Concluding Comments

This chapter has overthrown several long-standing presumptions concerning trade, growth, and government policy. We have shown that, in a learning economy, free trade is not, in general, desirable. Growth—and societal welfare—is maximized with some trade intervention to encourage the industrial (learning) sector. While we have criticized the infant-industry argument for protection, we have provided an alternative, more general argument—the infant-economy argument for protection. Intervention can be desirable, *even if the infant never grows up*.

Constraints on the set of interventions, such as those arising out of international trade agreements, may be welfare reducing. In response, it may be optimal for governments to intervene in the exchange rate—to such an extent that it may be desirable for them to run surpluses.[22]

With these interventions, developing countries may be able to close the knowledge gap that separates them from more-developed countries. Depending on the nature of the spillovers from the developed to developing countries, eventually, the knowledge gap may be closed, or, more typically, there may be a long-run equilibrium in which there is no convergence—there is a persistence in the relative gap. There is some evidence that, in many cases, gaps in per capita income have persisted—far more than one would have expected, say, from the Solow growth model, which had predicted convergence.[23] Such growth patterns are consistent with the analysis presented in this chapter. (See also Appendix B.)

There is a possibility of multiple equilibria—a developing country can get trapped in an equilibrium with a low level of per capita income (large knowledge gap) relative to the more developed country—a gap so large that investments in learning have a low return. When there are such multiple equilbria, history matters. The policy of structural adjustment foisted on African countries in the latter part of the twentieth century, which led to the deindustrialization of sub-Saharan Africa, may have contributed to some countries in that region being trapped today in such a low equilibrium.[24]

Openness and Learning

One of the reasons that our analysis differs from the conventional wisdom that emphasizes the virtues of trade is that it is often assumed that the best way that countries learn is through trade. Since trade barriers restrict trade, they restrict learning. We have, by contrast, highlighted the role of production (and investment) within a country. Properly formulated, the two hypotheses need not be in contradiction. We are not arguing for autarky. Countries that cut themselves off from the rest of the world are obviously cutting themselves off from important learning opportunities, and these countries will not perform well.[25]

Learning can be a function both of domestic production in general and of exports in particular. Having demanding buyers and competing in a global marketplace can facilitate learning. In this perspective, though, it is not trade in general, and not imports, that generate learning. Thus, trade liberalization that destroys a country's domestic production will destroy its learning in these sectors, with adverse effects on societal learning, if there are significant learning spillovers from these sectors to the rest of the economy. As we argued in this chapter, if trade liberalization leads the economy to specialize in sectors where there is little learning, the economy will not grow as well as it would with protection. What matters is not trade, but what is traded, and in particular, what is produced and exported.

Being embedded in a global economy has distinct advantages, but there are also risks and costs. Successful economies have figured out how to manage globalization.[26] For instance, joining the WTO opens up markets for a country's goods, but restrains the country's ability to pursue industrial-learning policies; China and other East Asian countries developed other tools (such as the use of exchange rates) as alternatives. The United States and other advanced countries have tried to foist an IPR regime, which is designed to serve the interests of the advanced countries (or more particularly, certain special interests within those countries) on the rest of the world. India and some other emerging markets have been better at shaping their own IPR regime, within the confines of the strictures imposed by TRIPS (the *trade-related intellectual property regime* of the WTO).

There are also ideas that disseminate from the advanced countries to the emerging markets and developing countries that may or may not lead to better economic performance. The more successful countries have done a better job at discriminating among these ideas. The Washington Consensus's ideas about good development policies have sometimes been forced on countries as a condition for their getting assistance. But some countries have adopted these ideas as their own; in effect, they have uncritically accepted the models on which those policies are based, not fully realizing that while those models provide a poor basis for understanding advanced countries, they are even less well suited for developing economies. Fortunately, since the 2008 crisis, those models, and the policy advice based on them (e.g., financial and capital market deregulation), have been looked at more critically.

In the following paragraphs, we describe other broader implications of the analysis of this chapter.

Global Imbalances, Excess Reserves, and Global Perspectives

Some countries have been criticized for contributing to global imbalances by accumulating excessive reserves. In static models, it has seemed irrational for developing countries—suffering from capital shortages and with constrained consumption—to do so; just as it has seemed peculiar that the United States, with an aging population, is running long-term deficits. This chapter shows, however, that once dynamic learning benefits are taken into account, with sufficient constraints on industrial policies, such as those imposed by the World Trade Organization, the accumulation of reserves by a developing country, beyond a level required for precautionary reasons to manage global volatility may be reasonable if the learning benefits are large enough. With these restrictions, governments are forced to undertake second-best measures, such as the exchange-rate interventions described earlier. It is perhaps not an accident that China's reserves and surpluses soared after it joined the WTO, as it faced new restrictions on its industrial policies simultaneously with new opportunities to export.

The effect of these interventions is almost surely to lower growth from what it otherwise would be—reducing the pace at which the gap between developing and developed countries is closed. They also increase the welfare costs of intervention, since they reduce the set of interventions that are permissible.

Interestingly, while these policies lead to a lower level of imports initially, because of the induced growth, over the longer run, the country's level of imports is actually increased.

It is obviously impossible for all countries to lower their exchange rates or expand their exports of manufactured goods through subsidies. It is impossible for all countries to use these policies to advantage themselves relative to others, even if it is possible for one small country—or even several of them or a large country like China—to do so. Of course, as the analysis of chapter 7 made clear, from a learning perspective, it *is* desirable for all countries to expand manufacturing (or other learning sectors) relative to other sectors, so policies that do encourage learning sectors (including subsidies) *are* desirable, even from a global perspective.

This may be true even if some of the increases in manufacturing in the emerging markets and developing countries come at the expense of production of those goods in the developed countries. The gains in learning in the former may more than offset the losses in the latter because the emerging markets and developing countries have so much more to learn as they strive to catch up. This is perhaps one of the reasons that global growth in the last thirty years has been so strong. It is not the expansion of trade as such. Because it was based on trade in commodities, from which there were few learning benefits, pre-World War I expansion of trade did not lead to such growth. The developing countries were exporting, but not growing. It is not exports in general that lead to growth, but exports (or more precisely, production) of manufactured goods and certain other products.[27]

What are the policies that maximize growth of the global economy—that facilitate learning and the transmission of knowledge *globally*? Just as, at the national level, the precepts that maximize static efficiency often run counter to those that maximize learning and growth, so too at the global level. WTO rules are not designed to maximize global growth. To the extent that they are based on economic theory (as opposed to power and special interest politics), it is a static theory that ignored endogenous learning. An international trade regime that sought to increase learning should presumably have more scope for developing countries to undertake industrial policies.

While this chapter has focused on learning's relationship to trade, learning is also affected by other aspects of globalization and the rules that govern it. Chapter 10 highlights the effects of the rules governing finance, how they may impede institutional learning about how to allocate capital and manage risks by financial institutions, how the allocation of capital by global capital markets does not take into account domestic learning benefits, and how existing rules and policy precepts have contributed to macroeconomic instability, which, as chapter 4 emphasized, is *not* conducive to learning.[28] The rules governing migration—the movement of people across boundaries—also has significant effects on learning; one of the important ways that learning is promoted is migration (through what are called cultural remittances), especially when it is temporary. On the other hand, permanent migration of a country's most talented young people may have a very negative effect on a country's ability to learn.

Changing Comparative Advantage

The central model that we explored in this chapter is one in which there are full spillovers, so that if the less-developed country initially has a comparative advantage in agriculture/crafts, it always does. That means that if the country wants to have an industrial sector, it must *permanently* provide some protection. It is perhaps incorrect to say that the infant never grows up: productivity in manufacturing may increase enormously, and the gap between productivity in that sector in the developed and developing country may narrow markedly. But because of the assumption of full spillovers, *comparative* advantage never changes. As we have repeatedly emphasized, even if it were true that infant industries sometimes never *fully* grow up, the support provided by the government to the industrial sector pays off: the economy is on a longer-term, faster growth trajectory than it would have been otherwise.

But Korea represents the more typical story, where, as it learns, productivity in the industrial sector increases faster than in agriculture, so much so that *eventually* the country's comparative advantage changes. That means that *eventually* government intervention to maintain a relatively large (larger than would be the case under unfettered market forces) industrial sector is no longer required. But as the discussion of the previous section points out, even after the country achieves some success in improving manufacturing capacities, it may still want to intervene, to produce *more* manufacturing goods than it otherwise would have produced.

Looking Forward

The analytic discussions of this book provide the foundations of the policy discussions of part 2. We have established the desirability of government intervention to encourage the industrial sector and to help create a learning economy and society. There are many ways this can be done and many policies that affect the economy's learning. There are many concerns in the design of such policies besides those upon which we have focused here. Some policies may be easier to administer; others may be more immune from capture by special interest groups. The following chapters focus on some of the most important of these policies. Not surprisingly, as we will explain, the most successful countries have, at critical junctures in their growth, undertaken these policies.

APPENDIX A

Trade Interventions and Optimal Tax Theory

It may be useful to set our analysis from the first section of this chapter into the context of optimal tax theory. The classic works of Diamond and Mirrlees (1971a, 1971b) showed that production efficiency was optimal. That implied that it was not desirable to interfere with trade efficiency. There should be no trade taxes.

Our general analysis differs from that classic theory by assuming that there is a social value to the production of manufactured goods. In Diamond and Mirrlees, all that society values is *consumption*. Once we recognize that there is a value to the domestic production of manufacturers, it follows that we want to distort production patterns to increase the output of manufacturing, reversing the Diamond-Mirrlees presumptions.

Once we introduce restrictions on the set of admissible taxes, we are in the third-best world explored by Dasgupta and Stiglitz (1971, 1972), which showed again that there was a presumption in favor of distorting production efficiency, including imposing trade interventions (Dasgupta and Stiglitz 1974; Blitzer, Dasgupta, and Stiglitz 1981). The basic insight of that literature is that one uses the set of admissible taxes to partially substitute for the restricted taxes. If we cannot directly encourage the production of industrial goods in a targeted way, we encourage the broader class of goods that we are allowed to promote, "tradables." To ensure that the demand and supply of nontradables is equilibrated, we have to introduce still further distortions in the economy, but (at least for small trade interventions) these have small welfare costs, e.g., consuming less today than we otherwise would, and consuming more tradables than we otherwise would. These changes in consumption can be induced either directly (as a result of consumption subsidies) or indirectly (through government expenditure programs directed at manufacturing goods).

APPENDIX B

Diagrammatic Exposition

In this appendix, we provide a simple diagrammatic exposition of the analysis of this chapter, focusing on a model with two goods (A and M), in which learning occurs in the manufactured goods but spills over fully to the other sector. Each good is produced by using a constant amount of labor per unit output. Learning reduces the labor requirement.

Short Run Equilibrium

Figure 8.1 shows the production possibilities schedule of the two countries (on a per person basis), with the developed country being able to produce (per person) more of both goods (i.e., it has an absolute advantage in both goods, but a comparative advantage in manufacturing).

Figure 8.2 shows the consumption possibilities curve of the developing country, if it specializes in agriculture. By specializing in agriculture, its "consumption possibilities schedule" is unambiguously improved. Under free trade, it will do this, choosing the point along the consumption possibilities curve that maximizes its utility.

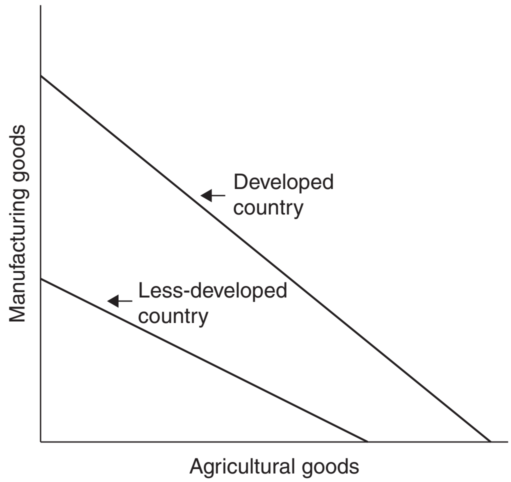

Figure 8.1 Comparative and absolute advantage.

The developing country has an absolute disadvantage in both goods, but a comparative advantage in agriculture.

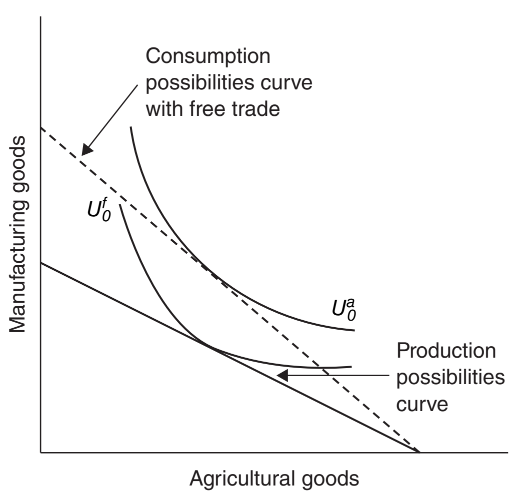

Figure 8.2 Free-trade equilibrium.

By specializing in agriculture, the developing country could have a higher level of utility—more consumption of both goods. This is the free-trade equilibrium.

The result, as noted in the text, is that over time the developing country's production possibilities schedule is unchanged from the previous period, while the developed country's production possibilities schedule has shifted out, but in parallel, because of the learning and perfect learning spillovers. Because the relative price in the large developed country remains the same, the developing country's consumption possibilities schedule remains the same. The developing country continues to specialize in agriculture goods and continues to stagnate.

Figure 8.3 shows that under autarky, in the short run, consumers are worse off. Moving from free trade to autarky lowers utility from U_0^f to U_0^a, where the subscript 0 refers to the period and the superscript refers to the trade regime: *f* for free trade and *a* for autarky.

But because with autarky, Korea (our proto-typical developing country) is producing manufactured goods—and learning—its production possibilities schedule is moving out. At some later date, T, it has moved out so much that society is better off (even with the trade distortion) than it would have been had it remained with free trade. With free trade, there would have been no learning, so $U_T^f = U_0$ where the subscript T refers to utility at time T. Thus, U_T^f is less than U_T^a.

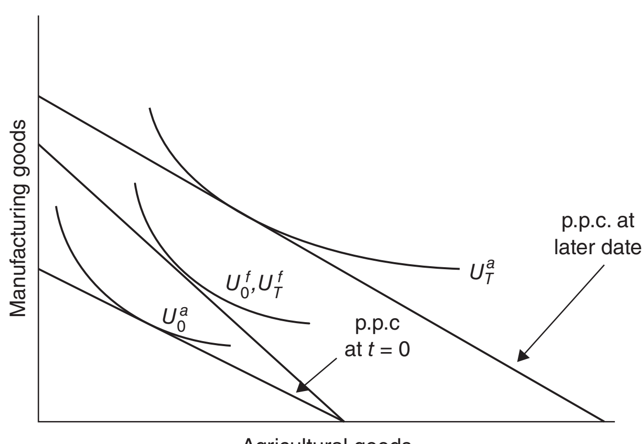

Figure 8.3 Quota autarky.

With autarky, in the short run, welfare is lowered: $U_0^a < U_0^f$. But in the long run, it is increased: $U_T^a > U_T^f = U_0^f$.

Optimal Trade Interventions

In figure 8.4, we model the trade intervention as an import quota on manufactured goods, which means that the residual between domestic demand (at the international price) and domestic production is met by domestic production. Thus, a quota translates into moving the "consumption possibilities" schedule from AM′, downward to AM″, and the equilibrium consumption moves down from E_0^f, the free-trade equilibrium, to E_0^q, the equilibrium with a quota. Utility in the short run is clearly lower, and the larger the trade restriction, the lower welfare is. The limiting case is that of autarky, discussed earlier.

With a quota, the static distortion is lower than in autarky, and the growth benefits are smaller. But utility in the long run, U_T^q, is still higher than with free trade.

The optimal quota balances the marginal loss of utility in period t from restricting manufactured goods a little more with the marginal gain in utility in period T from the extra learning.

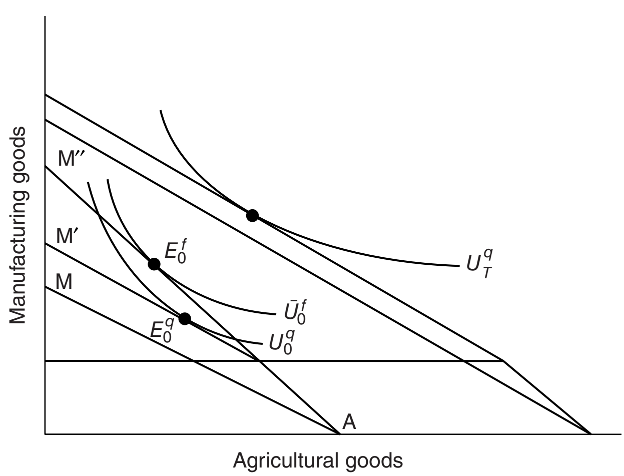

Figure 8.4 A quota.

With a quota, the static distortion is lower than in autarky, and the growth benefits are smaller. But utility in the long run, U_T^q, is still higher than with free trade.

APPENDIX C

Equilibrium with Leaders and Followers[29]

In the text, we argued that there can be an equilibrium in which some countries persistently remain leaders, and others followers. Here, we provide a simple graphical analysis of that equilibrium.

This analysis turns on its head a central contention of neoclassical growth theory since the work of Solow (1956), arguing that countries with different initial conditions should converge. In fact, the evidence on convergence has been disappointing.[30] Our theory explains this absence of convergence.

In the leader-follower equilibrium, it is optimal for laggard countries to remain laggards and never catch up. They sufficiently benefit from the dissemination of knowledge from the leader that it doesn't make sense for them to make the "Big Push" to join the club of leaders. But, except in some limiting cases, the followers are not fully passive—they pursue policies designed to close the gap between themselves and the leader, but even as they do so, the leader pursues policies that open the gap further. Not surprisingly, even though both leaders and followers pursue "innovation" policies, the policies that are optimal for each can be markedly different.

We should emphasize at the onset that while we talk about leaders and followers, our characterization is too stark. Knowledge is multidimensional. Some firm/country could be on the knowledge frontier along some dimension, but well within the frontier on another. That is certainly true among countries that claim to be "at" or "near" the frontier, implying that they have a considerable amount to learn from each other.

The central hypothesis is that the growth rate of the developing country is a function of the (share of the) labor force allocated to manufacturing (denoted by π^L) and the gap in productivity κ, where κ is the ratio of the output per worker in manufacturing in the developed country to that in the less developed country (so $\kappa > 1$).[31]

The crucial assumptions that distinguish this model from conventional growth theory are that productivity growth is *endogenous* and that knowledge does not flow freely across borders. In conventional growth theory, the rate of growth is exogenous, not affected by anything the firm (or society) does, and knowledge flows freely, so $\kappa = 1$. While some forms of knowledge do move easily across borders, many others (tacit knowledge, knowledge related to the conduct of particular institutions) may be far less mobile than labor or capital.

As we noted in the text, as the developing country allocates more labor to manufacturing, its short run welfare decreases, but its long run growth increases. As a result, there is an optimal value of π^L for each value of κ. Since there is more to learn when κ is smaller (the gap in knowledge is larger), we assume that when κ is smaller, a larger fraction of labor is allocated to manufacturing. In figure 8.5 this is depicted as the downward sloping curve *PM*.

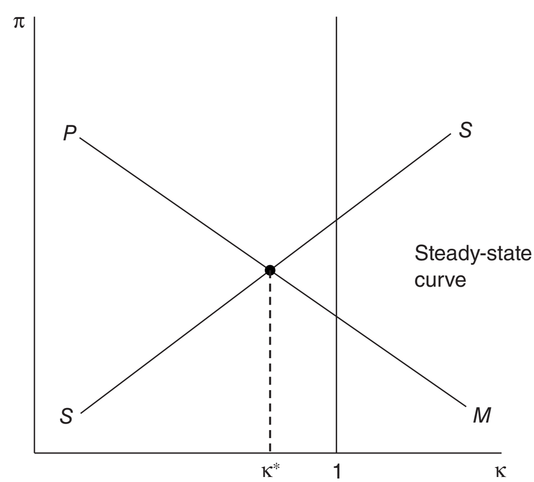

Figure 8.5 Catching up: steady-state equilibrium.

The steady-state equilibrium, showing the long run allocation of labor to manufacturing and the steady gap between the leader and the follower. Along the steady-state locus, *SS*, as κ increases, the pace of learning slows (there is less to learn) at any value of π^L, and so for the developing country to maintain the same distance from the frontier, π^L must be increased. The steady-state equilibrium entails the catching-up firm remaining always behind. The *PM* locus shows the profit maximizing value of π^L for any value of the gap. As the gap increases, we assume that not only does the level of learning at any π^L increase, but so does the marginal return to "learning by doing," so that π^L increases.

In this simple model, what the small developing country does has a negligible effect on the developed country; it continues to grow at the rate g^d. But for there to be a steady gap between the two, the developing country must have the same rate of productivity increase. At any given gap, the greater the allocation of labor to manufacturing, the faster the growth. It is natural to assume that when the gap is larger, at a fixed allocation of labor, there will be a faster pace of productivity growth—simply because there is more to learn, there will be more learning. This means that for productivity in the developing country to grow at the same rate as in the developed country, as κ increases, so must π^L, illustrated by the *SS* curve in figure 8.5.[32]

Figure 8.5 shows the "normal" case where there is a unique leader-follower equilibrium.

The infant never fully grows up, but to keep up with big brother, he has to continue to have industrial protection. There is a benefit to being the laggard: it is able to maintain the same rate of growth of the developed country by taking advantage of the knowledge that flows down from the developed country with a small fraction of its labor force allocated to the industrial sector. It can take *some* advantage of its comparative advantage in agriculture.

Other possible configurations may emerge: the infant may catch up; it may be optimal to have no industrial policy and simply absorb whatever knowledge trickles down to it; and there can be multiple equilibria—countries can be trapped in a low level equilibrium marked by a high knowledge gap, but with a positive enough boost, those countries can move into a better steady-state equilibrium, with higher levels of consumption and a much smaller gap with the leading countries. (See Stiglitz [2014e] for the conditions associated with each of these possibilities.)

Impact of Industrial Policies in Advanced Countries

Figure 8.6 shows the impact of industrial policies in advanced countries. If the advanced country pursues an industrial policy, g^* will be higher than it otherwise would have been, so π^L has to increase at each κ (i.e., the *SS* curve, defined by [8], shifts up).

The steady-state equilibrium depends on whether the advanced country pursues an industrial policy (i.e., takes into account that its

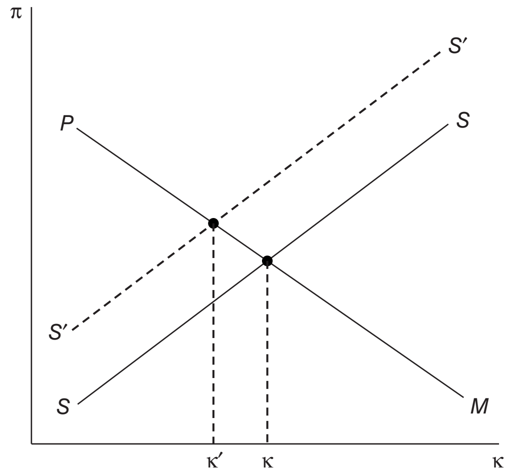

Figure 8.6 Catching up: Multiple steady-state equilibria.

If the gap between the advanced and developing country is large, there can be little learning by the developing country, and it can get caught in a low-level equilibrium trap.

growth rate g can be affected by its labor allocation). If it does (and our previous analysis showed that normally it would want to do so), g^* will be higher than it otherwise would have been, so π^L has to increase at each κ; in other words, the *SS* curve shifts up. This in turn means that (a) in steady state, the developing country will also have a higher growth rate; but (b) it will have to have a stronger industrial policy (i.e., a greater distortion in the static allocation of labor); and (c) the equilibrium gap between the developed country and the less developed country will be larger. (See figure 8.6.)

CHAPTER NINE

The Role of Industrial and Trade Policy in Creating a Learning Society

INDUSTRIAL POLICIES—meaning policies by which governments attempt to shape the structure of the economy, including the choice of technique and the sectoral allocation of the economy—are back in fashion, and rightly so. The major insight of welfare economics of the past fifty years is that markets by themselves in general do not result in (constrained) Pareto efficient outcomes (Greenwald and Stiglitz 1986).

By now, there is a rich catalog of market failures, circumstances in which the markets may produce too little of some commodity or another, too much of another, or may result in too little employment, and in which industrial policies, appropriately designed, may improve matters.

This chapter focuses on one of the central reasons for industrial policies, one that has been at the center of this book: Markets on their own do not create a learning society; the structure of the economy that results from market forces results in less learning—and less growth—than there could or should be.

We believe that one of the objectives of economic policy should be to create an environment that enhances both learning and learning spillovers. As we argued in chapter 1, creating a learning society is more likely to increase standards of living than small, one-time improvements in economic efficiency or improvements that derive from the sacrifices of consumption today to deepen capital.

This chapter moves from the broad analytic discussion of the previous two chapters to the central policy debates. Industrial policies have been highly controversial, and trade policies, including those designed to help restructure the economy, even more so. We explain why many of the arguments used against these interventions are misplaced and suggest how trade and industrial policies might more effectively contribute to creating a learning economy.

The chapter is divided into eight sections. In the first, we explain why much of the debate about industrial policies is misplaced—whether governments want to or not, they are, in practice, always engaged in industrial policies. In the second, we focus more narrowly on developing countries, arguing that these policies are especially relevant to such countries. In doing so, we refute the long-standing Washington Consensus presumption *against* industrial policies. In the third, we turn to the objectives of industrial policy. This is followed by a more extensive discussion of trade policy as an instrument of industrial policy, a continuation of the analysis of the previous chapter. The final three sections of the chapter contain more general reflections on industrial policy—its historical role, the role of political economy, and strategic considerations.

1. The Inevitability of Industrial Policies

Governments are inevitably involved in industrial policy, in shaping the economy, both by what they do and do not do. If they don't manage the macro-economy well, then more cyclically sensitive industries will be discouraged. If they use interest rate adjustments to stabilize the economy, interest-sensitive sectors will suffer. If they don't stabilize the exchange rate, then nontraded sectors are encouraged.

Moreover, in almost all countries, governments play a central role in education, health, infrastructure, and technology; and policies and expenditures in each of these areas—and the balance of spending among these areas—also shapes the economy.

Markets don't exist in a vacuum, and each of the laws and regulations that structure our markets—such as laws governing bankruptcy and corporate governance—shape the economy. Development economics routinely emphasizes as central to growth the study of institutions. All the rules and regulations, the legal frameworks and how they are enforced, affect the structure of the economy. Therefore, unwittingly, government is always engaged in industrial policy. American laws giving priority to derivatives in the event of bankruptcy was an industrial policy that encouraged derivatives. American laws providing that student loans cannot be discharged, even in bankruptcy, discouraged the education sector. Tax systems that tax financial speculation more lightly than other forms of economic activity encourage resources to move into financial speculation.

In short, all governments have an industrial policy, explicit or otherwise. The only difference is between those who construct their industrial policy consciously and those who let it be shaped by others, typically special interests, who vie with each other for hidden and open subsidies, for rules and regulations that favor them over others.

Even the agenda of financial market liberalization was an industrial policy—one pushed by the banks and the financial sector, the effect of which was to lead in many countries to a bloated financial sector, rife with explicit and implicit subsidies (reaching record levels in the crisis of 2008–2009), diverting resources from other uses that arguably would have led to higher sustained growth. It was an industrial policy that led to more macroeconomic instability, which, as we explained in chapter 4, itself adversely affected learning.

The Instruments of Industrial Policy

We have explained how government, in all of its policies, laws, and regulations, needs to be mindful of their effects on the structure of the economy and on the creation of a learning society. Earlier chapters explained why government intervention to promote learning was desirable. But there are many forms this intervention can take, many actions by which the government can help shape the economy. Since we have argued that virtually all actions undertaken by the government have some effects in shaping the economy, our analysis has to be confined to those where the impacts are largest, or which are explicitly directed at shaping the economy. This and the following four chapters look at several of the key policies: This chapter at industrial and trade policies, the next at financial policies, chapter 11 at macroeconomic policies (including exchange rate policies) and investment policies, and chapter 12 at intellectual property.

2. The Special Importance of Industrial Policies for Developing Countries

Closing the Knowledge Gap

Earlier, we emphasized that what separates developed from developing countries is not just a gap in resources but a gap in knowledge (Stiglitz 1999b, World Bank 1999). Much of the difference in per capita income between these countries and the more advanced is attributable to differences in knowledge. If this is so, then development strategies should be centered on promoting learning, closing the knowledge gap between more- and less-developed countries. Policies that transformed their economies and societies into "learning societies" would enable them to close the gap more rapidly, with marked increases in incomes.[1] Development entails learning how to learn (Stiglitz 1987c). As we argued earlier, too, the fact that some countries and firms have "learned how to learn" helps explain why the last two centuries have seen such remarkable increases in standards of living, in comparison to the millennia that preceded them, which were marked by stagnation.

How Structural Adjustment Policies Stifled Growth

But rather than promoting the learning sectors, the policies foisted on developing countries by, say, the international economic institutions, have actually discouraged the learning sector (industry) in many developing countries, especially in Africa (see, e.g., Noman and Stiglitz 2012a, 2012b, 2012c; and the references cited there). The result is that over the past thirty years, Africa has suffered from deindustrialization. The quarter century from the early 1980s was a period of declining per capita income and increasing poverty.

Structural adjustment policies advocated by the IMF and the World Bank were predicated on the belief that by eliminating "distortions" in the economy, Africa would grow faster—by constructing an economy based on principles of free and unfettered markets. The reigning dogma was that with the government restrained to ensuring macro-stability, which typically just meant price stability, economic performance would increase and all would benefit. In focusing on static efficiency, these international institutions totally ignored learning and the associated dynamics.

But this was not the only way that the structural adjustment programs impeded the development of a learning economy. Much learning occurs on the job—but if there is to be on-the-job learning, there has to be jobs. It was recognized, of course, that eliminating trade protection would result in the loss of jobs, some in agriculture, many others in industry. The strongly held belief, however, was that these workers would quickly find jobs in new industries, consistent with the country's comparative advantage. Moving resources from inefficient protected sectors to more efficient competitive sectors would raise incomes.

Things didn't turn out as the advocates of these policies had hoped. Rather than growth, there was decline. Job creation often (or even typically) didn't keep pace with job destruction, and so workers moved from low-productivity protected sectors to even-lower-productivity unemployment, open or disguised.

The Washington Consensus and Learning

The structural adjustment policies in Africa were associated with a set of economic doctrines that shaped the policies demanded of developing countries if they were to receive assistance from the West. For more than a quarter century, policies, especially in developing countries, were dominated by a set of ideas that is commonly called the Washington Consensus. It was these ideas that led to the structural adjustment policies that in turn led to Africa's deindustrialization,[2] with such adverse consequences for its growth and the well-being of its citizens.

The Washington Consensus policies were predicated on the assumption that markets, by themselves, are efficient, and that therefore the major source of inefficiency or mal performance of the economy arises from government intervention. Hence, the first item in the reform agenda is to eliminate these interventions with the market. The only (or at least the main) economic role of the government was to ensure price stability and property rights (including the enforcement of contracts).[3]

Thus, the Washington Consensus, and the ideology on which it was based, gave short shrift to market failures. When they grudgingly admitted to market failures, they suggested that government was not capable of correcting these market failures, because of "political economy" reasons.

This chapter explains why both parts of these doctrines were wrong: Market failures are pervasive, and governments—even in developing countries—can improve matters, and have done so, even if they have not "perfectly" corrected them.

In their aversion to industrial policies, the Washington Consensus policies focused on static efficiency. They didn't even consider what the consequences for innovation and learning were. If there was learning and technological progress, it was assumed to be exogenous, outside the purview of policy, and certainly outside the purview of the economic policies on which they focused. This was so was striking, given the observation, made earlier, that development is so much about learning and economic transformation.

LEARNING AND ONE-SIZE-FITS-ALL POLICIES One critique of the Washington Consensus is that it has attempted to impose one-size-fits-all policies. Such policies may be particularly inappropriate when it comes to creating a learning society.

A critical aspect of "learning" is that it takes place locally and must adapt to local differences in culture and economic practice. Thus, "learning" prescriptions that work in some environments will not work in others. For example, in some economies, especially in East Asia, close relationships between government and business seem to have helped development; potential conflicts of interest have been contained, and there have been marked benefits from effective coordination (see World Bank 1993). But such relationships can easily evolve toward "crony capitalism," and the associated corruption may impede development.[4] Learning how to relate to government has value in most economies, but in some, the skills required may concern those related to bidding processes, in others to interpersonal connections. American firms have had to learn to adapt to the Foreign Corrupt Practices Act.[5] Labor norms differ too among countries, and personnel policies have to accommodate such differences. Differences in consumer preferences and norms as well as in distribution channels necessitate different "learning" about marketing. Most importantly, and perhaps obviously, relative factor prices may differ, so that the returns to learning how to save on the utilization of one factor versus another may differ.

These cross-country differences have numerous implications. It helps explain why learning in a firm may spill over more easily to other firms in the same country than to firms in other countries. The learning in one country may simply be less relevant to production in the other country.

They help explain too why it is that in some economies public enterprises function well, while in others they do not.[6] (Functioning well means, of course, being able to learn and adapt—and in countries that have more broadly created a learning society, or at least one where learning permeates an important segment of society, even public enterprises can learn and adapt.)

They also help explain the limitations of globalization: Local firms have a competitive advantage in having more knowledge about local circumstances (see Greenwald and Kahn 2009). Much financial information is chiefly available locally, and even when information is available, outsiders may have less of an understanding of the nuances of the country's distinctive institutional structure—as foreign investors have learned about U.S. mortgages. Thus, effective capital deployment will often require local financial institutions.

Unfortunately, Washington Consensus policies, which pushed capital and financial market liberalization, did not take into account the importance of this local knowledge. Foreign banks succeeded in attracting depositors away from local banks, because they were perceived as safer (and in some cases, may have been, because they had the implicit guarantee of governments with deeper pockets). But foreign banks were at an information disadvantage relative to local banks about small- and medium-sized local firms, and it was thus natural that lending be diverted away toward loans to government, consumers, and large domestic firms (including local monopolies and oligopolies). But in doing so, local learning and entrepreneurship may have been undermined and growth weakened.[7]

Industrial Policy and the Distinctive Circumstances of Developing Countries

There are several other reasons that in developing countries industrial policy should play an even more important role than it does in developed countries. The first two concern the fact that these countries are laying down the foundations that will provide their economic structures for decades to come.

We have emphasized how legal frameworks—the laws and regulations that govern a society and an economy—simultaneously *shape* the economy. They constitute a kind of industrial policy. But a key issue in development is the formulation of these laws, when they don't exist, and reforming the laws, often inherited as part of their colonial legacy, when they do. Developing countries have to be aware that the legal frameworks they choose are shaping their economy today—with important consequences for their future.

So too, the physical infrastructure that they invest in will increase the returns to some kinds of private investment and reduce that of others. Investments developing some ports or roads will lead to the development of the surrounding areas, at the expense of areas which might otherwise have been developed.

In both of these instances, the groups lobbying (for the institutional and physical infrastructure that serves their interests) are the firms that exist today; those that might exist with a different institutional and physical infrastructure have little or no voice.

Earlier, we noted that all countries have an industrial policy—but the industrial policy which is chosen by developed countries is chosen to advance their own economies, or special interests in their economies. Even if it were easy to borrow these ideas from the developed countries, and even if it is possible to design industrial policies that enhance the flow of knowledge from developed to developing countries, strengthening cross-border flows of knowledge should not be the only focus of a developing country's industrial policy. For instance, environmental impacts are important for all countries, but especially for developing countries. The fact that natural resources and the environment are "underpriced" means that there are insufficient incentives to allocate resources (including those devoted to learning) toward the environment and natural resources—so more resources get expended on saving labor.

This highlights a difference between developed and developing countries, and a reason why it is important that developing countries have their own innovation policies and an industrial policy which promotes indigenous learning. Much of innovation in advanced industrial economies has been directed toward saving labor. But in many developing countries, labor is in surplus, and unemployment is the problem. Labor-saving innovations exacerbate this key social challenge. (See chapter 6.)

Even when labor-saving innovation does not result in unemployment, it will have adverse distributional consequences, lowering wages. With inequality already so high in many developing countries, this should be of concern.

Behind this analysis are two general theoretical observations: First, when innovation is important, history matters, as the discussion of chapter 3 emphasized. Technology that is developed today affects the kinds of technological developments that it pays to make later.

Moreover, factor prices in developing countries differ from those in developed countries, and in both, factor prices differ from shadow prices, because of a variety of market failures.

These observations have, in turn, three important policy implications: In both developed and developing countries, industrial policies have to be strategic—to take into account not just the country's circumstances today, but its likely long-run situation. What matters is not static comparative advantage but dynamic comparative advantage. (What this entails will be discussed at greater length.) Second, both the direction of innovation which firms would undertake and that which governments should undertake can differ markedly between developed and developing countries. And third, in both developed and developing countries, governments need to shape the direction of innovation and learning. As we emphasized in chapter 6, economies have a choice of the direction of innovation—whether it is, say, resource saving or labor saving. Those choices should reflect scarcity values—what the relative shares of different factors would be if factor prices reflected true scarcity, i.e., if market prices and shadow prices were equal.

3. The Objectives of Industrial Policy

While this book focuses on the role of industrial policies (broadly understood, in the manner described earlier in this chapter) in promoting growth through learning, governments have to simultaneously be cognizant of other social and economic consequences. Industrial policy is usually conceived of as promoting growth, but it should be seen more broadly, as any policy redirecting an economy's sectoral allocation (or other production decisions, such as the choice of technique or the nature of innovation) where market incentives, as shaped by rules and regulations, are misaligned with public objectives. Governments are concerned about employment, distribution, and the environment in ways in which the market is often not. Thus, in those countries with persistent high levels of unemployment, it is clear that something is wrong with market processes: labor markets are not clearing. Whether the explanation has to do with inherent limitations in markets (e.g., imperfect information giving rise to efficiency wages[8]), unions, or government (e.g., minimum wages), the persistence of unemployment implies that "correcting" the underlying failures may not be easy. The social costs of unemployment can be very high, and it is appropriate for government to attempt to induce the economy to move toward more labor-intensive sectors or to use more labor-intensive processes.

Structural Transformation

It has increasingly been recognized that development requires the structural transformation of the economy (see Lin 2010, 2012; Lin and Monga, 2014; and Stiglitz 1998b). Markets themselves are not very good at such structural transformations, partly because the sectors that are being displaced—resources that have to move from one sector to another—typically suffer large capital and income losses and are thus not well placed to make the investments required for redeployment; and well-understood capital market imperfections (based on information asymmetries) limit access to outside resources (see Delli Gatti et al. 2012, 2013). One of the main impediments to the easy flow of resources from one sector to another is, of course, the lack of appropriate skills; a greater focus on "learning" would have led to policies that would have enhanced the relevant capabilities of the workforce.

Inequality

Many developing countries have been marked by high levels of inequality.[9] Industrial policies can affect the extent of inequality by increasing the demand for lower-skilled workers, driving up their wages, and lowering their level of unemployment. While policies focusing on distribution have traditionally been centered on tax and transfers, it may be better (more efficient) to have policies that change the before-tax-and-transfer distribution of income. Such policies reduce the burden imposed by distortionary redistributive policies (Stiglitz 1998b).

But there are further reasons that we should be concerned about growing inequality than just a sense of social justice. It can lead to increased political and social instability. There is, moreover, a growing understanding, even within the IMF, that inequality may lead to lower economic growth, more economic instability, and a weaker economy (see Stiglitz 2012c, 2011; Berg and Ostry 2011; and the references cited there). While there are many channels through which these adverse effects operate (e.g., inequality diminishes the aggregate demand for domestic nontraded goods, and central banks often take offsetting actions—lowering interest rates and relaxing regulations and their enforcement—which give rise to unsustainable bubbles), one may be of particular importance in developing countries, where there is a need for heavy public investments in infrastructure, education, and technology.

In a society with very little inequality, the only role of the state is to provide collective goods and correct market failures. When there are large inequalities, interests differ. Distributive battles inevitably rage, and to prevent redistribution, wealthy elites often try to circumscribe the powers of government. But in circumscribing government, they also restrict government's ability to perform positive roles. As we have argued here and elsewhere, government needs to play an important role in any economy, correcting pervasive market failures, but especially in the "creative economy."

Thus, our critique of noninclusive growth goes beyond that it is a waste of a country's most valuable resource—its human talent—to fail to ensure that everyone lives up to his or her abilities. Noninclusive growth can lead to democracies that do not support high-growth strategies. There can be a vicious circle, with more inequality leading to a more circumscribed government, leading in turn to more inequality and slower growth.

4. Trade Policy

Many of the industrial policies discussed in earlier parts of this book entailed government expenditures—subsidies to sectors which should be expanded or investments in R & D and education to enhance the competitiveness of the learning sectors. But governments in developing countries (and even in developed countries) have a hard time raising revenues (see, e.g., Aizenman and Jinjarak 2009). In the absence of lump sum taxes, all taxes are distortionary, and thus there is a real cost to providing such subsidies. (Our analysis in the preceding chapters took those costs into account in designing the optimal interventions.)

It is thus natural that governments try to shape their economy using tools that raise revenues rather than use them. Tariffs and auctioned quotas can do that (see Dasgupta and Stiglitz 1971, 1972, 1974; Emran and Stiglitz 2005). This helps explain the widespread use of trade policy as part of industrial policy.

But turning to trade policy encounters several problems. First, there is the widespread theoretical presumption in favor of free trade—dating back to Adam Smith and his attack against mercantalism; David Ricardo, who developed the theory of comparative advantage; and nineteenth-century liberal economists such as John Stuart Mill. Samuelson (1938) formalized the welfare benefits of free trade. And neoliberal doctrines of the latter part of the twentieth century, discussed earlier in this chapter, reinforced this presumption.

But the stringent assumptions required to prove the result that free trade was desirable only highlighted the limited scope of the result. The analysis assumed perfect markets—perfect competition, perfect risk markets, full employment, and, most important from our perspectives, no externalities and, especially from the perspective of this book, no learning externalities. When any of these assumptions is dropped, free trade is *not* desirable.[10] Chapter 8 showed that in a learning economy, growth and long-term welfare could be enhanced with trade interventions. The result that there should be no tariffs vanished even in the absence of limitations on government's ability to raise revenues. Thus, advances in modern economics have reversed the earlier presumption: Now there is a presumption *against* free trade. And this presumption is even increased when account is taken of possible adverse distributive consequences. Trade liberalization has been associated with increasing inequality.[11]

Second, whatever the theoretical presumptions, there is a widespread belief that, empirically, trade is good for growth. While some earlier cross-country regressions seem to confirm this politically popular view, upon closer examination, the evidence does not support the conclusion.[12] It appears that once the "misattribution of macroeconomic phenomena (overvalued currencies or macroeconomic instability) or geographical location (in the tropical zone) to trade policies… are corrected, any meaningful cross-country relationship between trade barriers and economic growth evaporates" (Helleiner 1994).

That trade liberalization *by itself* would not ensure growth should have been obvious from the large disparities that exist within developed countries: There are no trade barriers between northern and southern Italy (no barriers even to the movement of capital), and yet there have been persistent large differences in income. So too for the United States, until the federal government undertook actions (including assistance) that narrowed (but far from eliminated) the gap in income between the North and the South.

There is another way of seeing that trade (and trade liberalization) *cannot* be at the center of the explanation of successful growth: Trading opportunities available through globalization are universal; yet growth has been highly particular, both across countries (even among those that have liberalized) and within individual countries over time. It is particular local conditions that determine whether the universal trade opportunities lead to growth.

Much of the presumption that openness leads to growth is based on the observed correlation between growth and trade.[13] But that correlation does not prove causality: Is it trade that has caused growth, or vice versa? One way to get insight into this issue is to look not at trade, but at trade liberalization. Free trade is about trade liberalization. Developing countries that have grown more may have done so because they've exported more, not because they liberalized; they may have succeeded in exporting more because of enhanced learning, which could have been impeded by liberalization. The evidence supports this hypothesis—that for the most part causality goes from growth to trade. The United Nations Development Programme (UNDP 2003), among others, presents evidence that trade liberalization was *not* associated with higher economic growth. Similarly, the marked reductions in trade barriers facing the least-developed countries (under Europe's "Everything But Arms" initiative or the United States' AGOA) have not had the hoped for benefits in terms of an expansion in trade.[14]

This book (focusing on learning) helps explain why we should not be surprised at the result that trade liberalization may be adverse to growth, and in section 6 of this chapter, we present evidence that is consistent with these perspectives.

The third objection to trade interventions (as to other forms of industrial policy) concerns *political economy*: Even if, theoretically, there is a presumption against free trade, even if the evidence that trade is good for growth is, at best, weak, and even if the most successful countries employed trade policy as an instrument for economic growth, today's less-developed countries, it is alleged, are incapable of using trade policy as an effective instrument. Opponents claim that, more likely than not, it will be abused. We will confront this objection later in this chapter.

The final set of difficulties in the use of trade policies as an instrument of industrial policy arises from international trade agreements. In recent years, the WTO has attempted to restrict the use of such policies. These WTO restrictions on industrial policies and domestic sourcing (and possibly other restrictions on financial markets) may impede the ability of developing countries to foster learning, to garner for themselves the full learning benefits of foreign direct investment, or force them to employ second-best methods for promoting learning within their economies. Though the advocates of these restrictions surround them with neoliberal ideology, arguing that these restrictions are best for the countries themselves, the hypocrisy of countries like the United States and many in Europe attempting to restrict others from undertaking policies that were at the center of their own development has not gone unnoticed.[15]

The central result of earlier chapters was that while targeted subsidies may be the most effective way of promoting development and learning, trade policy (implemented through tariffs, quotas, or exchange rate policies) may be an effective instrument as well, as the last chapter demonstrated.[16] Though there are static costs associated with such interventions, the dynamic benefits can well exceed the costs. Chapter 8 showed that even though WTO constraints may impose a high cost on developing countries, trade intervention through managing exchange rates is still desirable, and later chapters will illustrate that there are other policies that can help shape the economy that can escape WTO strictures.

While trade policy can be, and has been, an important instrument for creating a learning economy, the learning perspective also can strengthen the rationale for regional integration: It may well be that individual national markets may be too small to support robust local industries. In that case the natural extension of the basic policy is to combine local economies at similar stages of industrial development into free-trade areas which are then protected by common uniform external industrial tariffs. Such extension has the added benefit of enhancing local competition.

Given space limitations, we cannot comment in depth on the form and design of trade policy. We make only two comments, elaborating in part on observations made earlier. First, our learning perspective emphasizes that what is important is not static comparative advantage but dynamic comparative advantage. As we explained in chapter 1, however, what that entails is more complicated than is often suggested. Indeed, we have argued that even the concept of static comparative advantage is more subtle than is widely recognized, once the (partial) mobility of capital, labor, and knowledge is taken into account. Three specific conclusions emerge from the discussions (including of chapter 1): (a) Countries cannot simply look to what other countries have done at similar stages of development. History matters. Changes in technology and the global marketplace mean that what worked for, say, Korea fifty years ago may not work for a country in a similar position today. (b) Given the difficulties of determining a country's (static or dynamic) comparative advantage, Lin's argument (2012, 2014) that countries should be careful about pursuing industrial policies that "defy" comparative advantage may not be of much help. (c) Countries have to be strategic—because history matters, they have to take into account how enhancing capabilities today (through trade policies) will affect their potential for capability enhancement in the future.

Second, more finely honed interventions require more information and may raise more political economy problems. Thus, while in the absence of these limitations, more finely honed interventions would be able to improve the dynamic-static trade-offs—getting more growth with a smaller sacrifice in current consumption/utility, taking them into account may imply the desirability of broader interventions. We explore this notion further in section 7.

5. History

There is ample evidence that many countries have successfully used trade and industrial policies. Indeed, there are few successful economies in which the government has not successfully employed industrial *and* trade policies, including trade restrictions. This is true not just for developing countries, like Korea, but also for developed countries like the United States.[17] Moreover, such policies can be used not only by countries with highly trained bureaucracies but by countries at early stages of development. At the time that many of the East Asian countries began their industrial policies, not only was their economic development lower than some of the less-developed countries today, but so too was their political development.

To understand the role of government in this arena, we need only think of the history of the United States, where the development of agriculture, the main "industry" in the nineteenth century, was promoted by government. The history of telecommunications—from the first telegraph line to the development of the Internet—entails one form of government support after another. Even the first browser was supported by the government. So too, in Brazil, we can think of the development of sugar-based ethanol and the Embraer airplane—and a host of other export-enhancing government interventions. In each of these cases, the social returns from these government-supported innovations are so high that they would support many less successful ventures.

The success of the most successful countries in development after World War II—those in East Asia—is largely attributable to their recognition of the importance of learning and the role of government in promoting it. Korea, for instance, paid little attention to its static comparative advantage. Its static comparative advantage would have led that country to focus on rice farming. But it knew that even if it became the most productive rice farming country in the world, its prospects would be limited. Only by focusing on sectors from which it could learn, and on the basis of which it could close the knowledge gap with more advanced countries could they achieve the growth they desired. Thus, Korea developed complementary industrial, education, and technology policies, and it succeeded, increasing its per capita income more than eightfold in a span of less than four decades.

Had it followed the dictates of the Washington Consensus policies, it would have eschewed industrial policies and focused investments in education at the primary level—and it would have, at best, been a middle-income rice-growing country. History has shown that not only have countries like Korea that adopted industrial policies done well, but also that those that have not have suffered; this is illustrated by the impact of the structural adjustment policies in Africa, which eschewed industrial policies, leading to deindustrialization and low growth.

Interestingly, Korea followed a strategy that was a mixture of the broad-based export promotion policies we have advocated as part of an infant-economy strategy and a more targeted approach (e.g., promoting chips and heavy industries and chemicals). Whether they would have done still better had they limited themselves only to broad-based strategies is an exercise in counterfactual history, for which the answers are at best debatable.

Historical Experiences with Trade Interventions

Historical evidence provides considerable support for the efficacy of learning-related trade restrictions.[18] While the most successful countries, both today (in East Asia) and historically (including the United States), have used a variety of instruments as part of their industrial policies, among the most important have been trade interventions. They have not only engaged in trade restrictions, but those restrictions were an explicit part of their growth strategies. Indeed, Rodrik (2001) argued that the three primary models of successful development in the twentieth century *all* relied on managed trade regimes: import substitution as practiced by a number of countries in the 1960s, outward oriented industrialization as practiced in East Asia in the 1980s, and the state-directed capitalism of China in the 1990s. Chang (2002) showed that almost all of today's rich countries used tariff protection and subsidies to develop their industries, and "Britain and the USA, the two countries that are supposed to have reached the summit of the world economy through their free-market, free-trade policy, are actually the ones that had most aggressively used protection and subsidies."

Of course, all of these countries, including the East Asian tigers, did engage in trade—but it was promoting exports, expanding the learning sector, not trade liberalization as is usually understood, opening up domestic markets broadly to foreign imports. Trade interventions have sometimes not worked out well. They have at times been used as protectionist tools by special interests, rather than to redirect society's resources toward creating a learning society. But the history of successful interventions suggests that failure is not inevitable. Hopefully, countries will learn from the failures (and successes) of the past, so that the returns from future interventions will presumably be greater than those from past interventions, a point to which we return later in the chapter.

Not surprisingly, countries like Myanmar, that pursued xenophobic policies that cut themselves off from others—both from trade and from learning—did not fare well.

In short, our analysis suggests that well-designed trade policies such as those in East Asia, centered on learning and the acquisition of technology, have done a far better job of promoting growth and learning than either of the extreme policies of full liberalization or full autarky.

War Time

Perhaps even more telling concerning the relationship between trade, growth, and learning is the fact that war times, in which trade is interrupted, have often seemed to be periods of enormous dynamic gains.[19]

Our learning perspective provides an interpretation of these seemingly anomalous experiences. Forced to rely on their own production capacities, economies increased their industrial production and, with it, their learning. Moreover, the exigencies of the moment "forced" the economies to learn more quickly than they might otherwise have done.

Does Latin America Demonstrate the Failure of Industrial Policies?

Some have argued that even if industrial policies were successful in East Asia, and even if they worked in times of war, they were a dismal failure in Latin America, and blame that continent's lost decade on its pursuit of industrial policies. Even if that conclusion were true, it would only imply that the form of industrial policy pursued in Latin America (import substitution versus the export-led growth of East Asia) was flawed, not that industrial policies per se were doomed to failure.[20]

But the conclusion that industrial policies were a failure in Latin America is, at best, contentious—at worst, simply wrong. Brazil, the most ardent adopter of such policies, had an impressive growth rate of almost 6 percent in the three-quarters century before 1980. Industrial policies played an important role in that country's success in this period. The lost decade was a result of Latin American countries' excessive indebtedness in the 1970s, the period of the oil shock—understandable, perhaps, given the low, or even negative, real interest rates at which the petro-dollars were being recycled—followed by an unprecedented increase in interest rates, a result of the United States suddenly switching its monetary policy regime to monetarism. The lost decade of the 1980s was, in short, a result of a macroeconomic shock, not a failure of microeconomic policies. The subsequent adoption of the Washington Consensus policies, which eschewed industrial policies, prolonged the subsequent period of slow growth. The more recent revival of growth, for example, in Brazil, has much to do with the government once again undertaking activist policies (Bértola and Ocampo 2012).

China and India

What about countries like China and India, which have liberalized and grown? A closer look at the timing shows that their take-off occurred prior to trade liberalization. Serious trade liberalization occurred *after* the trend rate of growth had increased. In both cases it was associated with "internal liberalization."[21] Reducing domestic distortions while maintaining external barriers provided precisely the conditions for the dynamic gains identified in this book.

History Shows That Industrial and Trade Policies Can Work

In short, the historical experience shows that industrial policies can work. They have worked in a variety of countries in a variety of circumstances, with a variety of strategies and instruments. Even instances of seeming failure need to be interpreted with caution. Good policies involve some risk—if every public or private investment succeeded, it would indicate insufficient risk taking.[22]

There are undoubtedly instances where industrial policy has failed because of abuses. But the relevant question is: Are the problems inherent in political processes? The next section shows that they are not, but that limitations in the political processes may affect the form that industrial policy should take.

6. Political Economy

A persistent criticism of industrial policies is that, even if market allocations are inefficient, even if market prices differ from shadow prices, government attempts to correct these failures will simply make matters worse. There is, it is alleged, just too much potential for misuse. Some go so far as to suggest that abuse is almost inevitable, because of a proclivity for rent seeking. Even if government doesn't *abuse* these policies, at least in developing countries it doesn't have the competency to implement them effectively.

Of course, most of these critics also believe that industrial policies are not needed: one should rely on the private sector. Government shouldn't be in the business of picking winners, of trying to outsmart the market. One should leave decisions about resource allocations to the market. As a former chairman of the U.S. Council of Economic Advisers is famously said to have claimed: It doesn't make any difference whether we produce potato chips or computer chips.

We have already provided answers to several of these perspectives. Market failures *are* rife. The objective of industrial policies is to correct these market failures.

That abuses are not inevitable is also clear: There is neither theory nor evidence in support of such a conclusion. The historical experiences cited in the previous paragraphs provide convincing evidence to the contrary. To be sure, there are instances of government failure, but none with the consequences or on the scale of the losses resulting, for example, from the failures of America's financial market before and during the Great Recession. As we noted, virtually every successful economy has employed industrial policies successfully, at one time or another (see Chang 2002, 2003; and the references cited there). And this is most notable in the case of East Asia.

But then the question is raised: Can there be effective industrial policies in countries with significant deficiencies in governance? The implication is that, while with "ideal government" intervention might improve matters, in the "real world—outside of a few isolated successes—interventions do not necessarily do so." The argument has been put that even if such policies contributed greatly to the success of East Asia, elsewhere they were less successful, because they were abused. These critics go on to say that, given the widely acknowledged deficiencies in governance in many African countries, they should shy away from such policies.

The critics point to Latin America and its lost decade as a prime example of the failure of industrial policies, but we have argued that Latin America's failure was a result not of its industrial policies but of its debt and macroeconomic policies. We have also pointed out that at the time Korea and some of the other East Asian countries undertook industrial policies, they were as economically and politically underdeveloped as many of the poor countries of the world today that are debating adopting such policies.

Framing the Question the Right Way

The strongest objection to these objections to industrial policy is that they frame the question the wrong way. The question is not whether in some cases such interventions have failed, but whether in some instances they have succeeded. The answer to that is unambiguously, yes. Indeed, the discussion in the previous section suggests that there are few, if any, successful countries that have not engaged in industrial policies.

Even with some failures, average returns have been positive. Indeed, again as we have already noted, good industrial policy entails risk taking, and with risk taking, we should expect some failures.

But even if average returns were sometimes, or often, low, the question today facing a country contemplating adopting industrial policies is whether it can learn from the successes and failures of the past, whether it can construct an industrial policy that has prospects of working.

Industrial Policies Are Not About Picking Winners

There is a second way in which the standard objections to industrial policy have framed the issue incorrectly. Industrial policies are not about picking winners. They are about correcting market failures in general, and creating a learning society in particular. It is now widely accepted that there can be important market failures arising from large negative externalities (e.g., from pollution or from excessive risk taking in the financial sector) and that there is an important role for government in correcting these market failures. Except among the extreme right, there is a general consensus that, even if government interventions may not have been perfect, we are a lot better off as a result of these interventions, which have curbed air and water pollution. We are concerned here with an equally important set of *positive* externalities, and we believe that government can be equally effective in (partially) correcting these externalities. Indeed, as the previous discussion has made clear, many governments (both in developed and developing countries) have a credible record of industrial policy interventions.[23]

Policies to Address Political Economy Problems

Moreover, these successes were not just accidents. The thrust of the book *The East Asian Miracle* (World Bank 1993; see also Stiglitz 1996;

Khan 2012; Stiglitz and Uy 1996) is that the East Asian countries used systemic procedures that limited the scope for abuse and increased the likelihood of identifying firms and sectors whose expansion would have society-wide benefits. First, the subsidies were limited—indeed, one of the main benefits government granted was access to credit. (With credit rationing, there is a difference between the shadow price of credit and the market price of credit; but the terms at which they got credit were still largely commercial.) Second, the government, to a large extent, employed contests for the allocation of the credit and other subsidies: those companies that demonstrated success in export markets gained preferential access to credit. It was not just profits that determined the allocation of credit; for higher profits might signal a better ability to exploit monopoly profits. But success in international markets demonstrated broader marketing and technical competencies. In effect, the government channeled rent seeking—the benefits of access to government subsidies and access to credit—making it socially constructive. Competition for rents led to firms that learned more and became more competitive in the global marketplace.

(In other countries, though, rent seeking has diverted resources away from growth-inducing activities, and even when resources are allocated to innovation, the innovation may not be growth inducing. Firms have devoted their resources to learning how to circumvent regulations designed to make the economy more stable and to learning how to exploit consumers and the firm's monopoly power better. Markets don't work well when private returns are not well aligned with social returns; and in those circumstances, incentives to innovate and learn are also distorted, as we have repeatedly noted.)

Governance and Institutional Reform

The fact that there have been some failures in industrial policy and some successes means that countries contemplating such policies need to think carefully about their design. There are institutional reforms that reduce the likelihood of abuses and increase the likelihood of successful interventions. East Asia provides examples of effective institutional designs that harnessed the drive for rent seeking in a socially constructive way. Other institutional innovations, including peer review, improvements in competitive bidding and allocational processes and transparency, time-limited programs, and the design of partnership arrangements (where those receiving assistance have to contribute significant amounts of equity) have contributed to reducing the scope for abuses and improving the efficacy of industrial policies. Some of these institutional improvements—including those related to the performance of development banks, such as Brazil's BNDES—show that there has been learning from the failures (and successes) of the past.

IMPLICATIONS OF GOVERNANCE DEFICIENCIES FOR THE DESIGN OF INDUSTRIAL POLICIES Reforming political processes is slow, and it may not be possible to make reforms fast enough to reduce the likelihood of abuses of certain forms of industrial policies to a tolerable level or to enhance the likelihood that they will be successful to the point where undertaking such policies seems advisable. The implication of deficiencies in governance is that *one needs to tailor the design of the instruments of industrial policy around the capabilities and governance of the public sector*.

This poses an important trade-off. Chapter 8 focused on the infant-economy argument for protection, centering on encouraging the industrial sector more broadly. An implication of the infant-economy argument is that trade restrictions should be broadly and uniformly applied to industrial products. Since the benefits sought are broadly rather than narrowly determined, there should be no attempt to support particular industries or, more precisely, to identify sectors in which there are larger spillovers. Broad-based measures such as exchange rate interventions require only that the government ascertain that the sectors that would be encouraged by such interventions, such as in the determination of the exchange rate, have more societal learning benefits than the sectors that would be discouraged—and there is ample evidence that that is the case (evidenced, in particular, by the success of export-led growth strategies). Firms and sectors within the economy self-select, and the expansion of firms and sectors with greater learning enhances the dynamism of the economy.

On the other hand, more targeted interventions can, if well designed and executed, lead to even more learning and faster rates of growth. Some countries have shown that they can manage the political economy problems of more targeted interventions. As we noted, the East Asian countries did so by using rule-based systems in which interventions were linked to past export success.

Ultimately, the test of the effectiveness of uniform infant-economy tariff policies is how well they have worked in practice, and here, at least superficially, the historical record is encouraging. The trade policy of the newly formed European Economic Community was, in the 1950s, one of high but relatively uniform external tariff barriers. The growth of the community behind these barriers was rapid. Similarly, Asian economies like Japan, Korea, China, Taiwan, and Singapore, while they did have some targeted interventions which, by and large, they managed well, have tended to favor broad rather than narrowly tailored barriers to trade, and they have all experienced strong growth. Finally, in its early history, the United States, too, tended to favor high and broadly applied industrial tariffs and succeeded in fostering high levels of growth.

Of course, there is no form of intervention that completely "solves" the political economy problem: Sectors that benefit from exchange rate intervention may lobby for the maintenance of that intervention even in the absence of learning benefits. Still, arguably, the political economy problems can be better managed with such broad-based interventions than with more narrowly directed interventions, which lead to the creation of narrowly focused special interests concerned with sustaining particular tariffs beyond their natural economic lifetimes. Properly designed, both the costs and the benefits of a uniform industrial tariff system should be widely dispersed. Moreover, a broadly based industrial tariff system should be, to some extent, naturally self-limited. Successful local industries should begin to export and, therefore, be naturally predisposed in favor of free trade.

A Methodological Response to the Political Economy Critique

There is also a methodological response to the political economy critique of industrial policies: Whether such political economy objections are true or not, the conclusion is based on political analysis, not economic analysis. And the political analysis is often more simplistic than economic analysis.

Moreover, similar questions can be raised about every other aspect of policy. Many governments have not used monetary and financial regulatory policy well; in some cases, the misuse can be traced to problems of governance. Many have argued that regulators and central banks in some advanced industrial countries were captured by special interests in the financial markets, and this played an important role in the 2008 global economic crisis (see, e.g., Stiglitz 2010b; Johnson and Kwak 2010). But this is not a reason for governments to eschew the use of monetary and financial regulatory policy.

LIBERALIZATION AND POLITICAL ECONOMY Finally, we note that liberalization is itself a political agenda. As we previously commented, markets do not exist in a vacuum. There are always going to be rules and regulations, even in a liberalized world. And the design of those rules and regulations will shape markets. The rules and regulations that were adopted in the process of "liberalizing" and deregulating financial markets in the United States and the U.K. led to bloated financial institutions backed by implicit guarantees from the monetary authority and ultimately the taxpayer—an industrial policy of deregulation and favorable regulation that distorted the economy.

7. Some General Reflections on Industrial Policy

Theory of the Second Best

Industrial policies distort consumption from what it otherwise would have been. Conventional economics (such as the Washington Consensus policies) emphasized the costs of these interventions. We have emphasized that when there are market failures (as is always the case when there are learning externalities), there will be benefits. Optimal policy weighs the benefits and costs as the margin.

The economics of the second best is particularly relevant here: R & D and learning give rise to market imperfections, sometimes referred to as distortions, where resources are not allocated in a "first best" way. Well-designed distortions in one market can partially offset distortions in others.

We use the word *distortions* with care: Common usage suggests that governments should simply do away with them. But as the term has come to be used, it simply refers to deviations from the way a classical model with, say, perfect information might function. Information is inherently imperfect, and these imperfections cannot be legislated away. Nor can the market power that arises from the returns to scale inherent in research be legislated away. That is why simultaneously endogenizing market structure and innovation is so important, as we noted in chapter 5. Similarly, the costs associated with R & D (or the "losses" associated with expanding production to "invest" in learning) cannot be ignored; they have to be paid for. Monopoly rents are one way of doing so, but—as we argue here—a far from ideal way.

As always in the modern economics of the public sector, the nature of the optimal interventions depends on the instruments and powers of government. Whether the government can abolish monopolies or undo their distortionary behavior has implications for the desirable levels of research and learning. It makes a difference, too, if the government can raise revenues to subsidize or support research or learning only through distortionary taxation rather than through lump sum taxes. But even when the government can only raise revenues through distortionary taxation, there are ways of doing so and spending the proceeds that increase societal well-being and the speed of innovation. (But the optimal investment in innovation is likely to be less with distortionary taxation than with lump sum taxation.)

Industrial Strategies

A key issue of industrial strategy is not only the direction (e.g., should Korea have attempted to reinforce its comparative advantage in rice, or to create a comparative advantage in some other area?) but also the size of the step. Should governments try to promote a nearby technology (product), nudging along a gradual, evolutionary process that might eventually have occurred anyway? Or should it take a big leap? The latter is riskier, bringing perhaps greater returns if successful, but also a higher probability of failure.

We have not formally modeled this critical decision, so the following remarks are only meant to be suggestive: The difficulties in learning increase significantly the bigger the leap; but so may the benefits. There are natural nonconvexities in the value of information/knowledge (Radner and Stiglitz 1984), implying that it pays to take at least a *moderate* step: small incrementalism is not likely to be optimal.

By the same token, using another analogy to corporate strategic policy, it pays to move to a part of the product space where there are rents which can be sustained (e.g., as a result of entry barriers, arising, for instance, out of returns to scale or specific knowledge). This almost surely entails not doing what others are or have been doing.

Among the most important insights of the theory of learning that we have developed is that, because history matters, decisions taken today concerning the direction of learning and innovation can have long-run consequences. Had Korea made different decisions in earlier decades, its economy today would look markedly different. The commodities which constitute its current static comparative advantage would be different, and its learning capabilities would be different. Because there are these long-term consequences, countries can't avoid making guesses about the evolution of the global economy. As we have argued, they can't or shouldn't simply follow patterns of development undertaken by countries at similar stages of development a quarter or half century ago.

Industrial policies are inherently risky; but it is even riskier not to have an industrial policy.

8. Concluding Remarks

Earlier, we argued that countries have no choice but to have industrial policies; budgetary policies and legal frameworks inevitably shape the economy. We have argued that countries need to be more conscious of the effects of policies in shaping the economy—and more active in doing so, with a particular focus on creating a learning and innovative economy. Markets don't do this on their own.

For developing countries in Africa and elsewhere, as they attempt to reindustrialize, to restructure their economies to become more integrated into the global economy and move away from excessive dependence on commodity exports, to raise standards of income, increase employment, reduce poverty and inequality, and to protect a fragile environment, industrial policies are of especial importance. We have explained why the widely cited objections—that though industrial policies may have worked in East Asia, they are inappropriate for Africa or other developing countries because of deficiencies in governance—are at best unpersuasive, at worst wrong. We have argued further that what is required is more than just creating a business-friendly environment, allowing the market to shape the economy on its own. Such policies are necessary, but they are not sufficient.

The debate today should not be about whether governments should pursue policies that shape the industrial structure of the economy.

Inevitably, they will and do. The debate today should center around the directions in which it should attempt to shape the economy and the best way of doing so, given a country's current institutions and how they will evolve—recognizing that the evolution of the institutions themselves will be affected by the industrial policies chosen. In the past, trade interventions have been an effective instrument of industrial policy; and although WTO strictures may limit the scope and form of trade interventions, they have not undermined the desirability of such interventions. Governance issues are, of course, relevant in all countries and are important in shaping the form that industrial policy takes and the instruments that are appropriately used. They strengthen the argument for broad-based interventions, such as those associated with exchange rate policy.

The belated recognition of the potential of these policies comes at a fortunate time, for changes in the global economy may afford the countries of Africa and some other regions that have lagged a distinct opportunity to transform their economies in a way that will, at long last, narrow the gap that separates their standards of living from that of much of the rest of the world.

This chapter is based partially on Greenwald and Stiglitz (2014a, 2014b) and Charlton and Stiglitz (2005, 2012), and borrows liberally from those papers.

CHAPTER TEN

Financial Policy and Creating a Learning Society

IF THERE are powerful arguments for broad barriers to imported industrial goods, those apply equally to restrictions on capital movements. Capital and financial services *within* a country can support learning; in contrast, financial services provided by foreigners can lead to a redirection of investment and learning *out of the country*, impeding the creation of a learning society. By the same token, the availability within a country of low-cost capital can encourage learning investments in the country. We saw in earlier chapters that complements to learning activities ought to be encouraged, and lowering the domestic cost of capital (by restricting capital outflows) may do so. This is especially so if learning is directly related to the level of investment, as Arrow originally hypothesized. Moreover, in a world of credit rationing, access to capital can be a key instrument of industrial policy.

There are further links between financial policy and learning. A poorly designed financial sector and poorly designed financial policies can lead to macro-instability, and as we noted in chapter 4, macro-instability impedes learning. (We will postpone the analysis of the links between financial policy and macro-stability to a later chapter.)

In short, a key objective of government policy should be to create a financial sector that fosters learning and helps create a learning economy. Standard policies (such as those advocated within the Washington Consensus) have simply ignored the effects of the financial sector on learning and have, as a result, fostered a financial sector which we believe is often not only not conducive to learning but inimical to it. This chapter will explore briefly the various links between financial sector policies and learning. What emerges from our analysis is a set of policy recommendations that is markedly different from those of the conventional wisdom. Underlying the analysis is the pervasiveness of externalities. While recent literature has emphasized the overall importance of externalities within the financial sector and between the financial sector and the real sector,[1] we emphasize here that private returns are often markedly different from social returns because of learning externalities.

The chapter is divided into four sections. The first focuses on financial market liberalization, opening up of countries to foreign financial institutions. The second centers on capital market liberalization, policies aimed at allowing the free flow of capital into and out of a country. The third takes up the issue of financial policy as an instrument of industrial policy—including how we can shape the financial sector in ways which encourage the development of learning sectors.

There is a certain symmetry between labor and capital. If creating a learning economy entails restrictions on the free flow of capital, it should be apparent that there are even more compelling arguments for shaping labor movements. In the concluding section of the chapter, we touch upon some key aspects of this complex issue.

1. Financial Market Liberalization

Information externalities are at the center of the debate over financial market liberalization. Western governments (directly, and through the international financial institutions) have strongly pushed developing countries to deregulate and liberalize their financial markets.

Deregulation ignored, of course, the market failures which had led in the first place to the demand for regulation of the financial sector. While a full discussion of these would take us beyond the confines of this book, suffice it to say that the actions of the financial sector can and do impose large externalities on the rest of the economy;[2] that there is a proclivity for the financial sector to engage in excessive risk taking; and that, as a result, economies with unregulated and underregulated markets are likely to have excessive volatility. This volatility imposes huge costs on society—most obviously on workers, but also on the public fisc, both directly, as it bails out banks in an attempt to prevent matters from getting worse, and indirectly, as the economic downturns have huge budgetary consequences, with declining revenues and increasing social expenditures.

Financial sector liberalization focuses on opening up a country's financial markets to banks and other financial institutions from abroad. It was one of the central initiatives of the Uruguay round trade negotiations. Ironically, even as the United States was pushing others to open up their financial markets, the United States had not yet even allowed interstate branch banking within its borders. Such restrictions (finally eliminated in 1994[3]) were introduced in the nineteenth century, because those outside the eastern United States worried that if the large New York banks were allowed into their markets, they would siphon off deposits, impeding the development of these other parts of the country—concerns closely related to justifiable worries of many in the developing world that have resisted full financial market liberalization.

One of the earlier justifications for such liberalization was that the domestic banks in developing countries would *learn* good practices from Western banks. In the aftermath of the global financial crisis, that argument seems less persuasive: They might have learned how to engage in more effective exploitation of the poor, through predatory lending and abusive credit card practices, and they might have learned about how to engage in deceptive off-balance-sheet accounting, but it is clear that there may be negative externalities associated with such socially unproductive learning.

The evidence presented in papers such as Rashid (2011); Yeyati and Micco (2007); Detragiache, Tressel, and Gupta (2008); Mian (2006); and Bayraktar and Wang (2004) is that financial market liberalization is not associated with faster growth in developing countries. A closer look at the evidence suggests some of the reasons (alluded to earlier) why this may be so.[4] Foreign banks lend less to small and medium-sized enterprises. And for good reason.

If an English bank opens a branch in Latin America, the managers know a lot more about investment opportunities in England than they do about investment opportunities in Latin America. Given the asymmetries of information, therefore, the natural tendency is going to be to take the financial resources and move them out of Latin America to England. And to the extent that it lends domestically, it will be in areas requiring less information—lending to domestic large enterprises, including monopolies, and to government.[5]

Just as the infant-economy argument argues for protection of the industrial sector, this analysis provides an analogous infant-economy argument for protection of finance. That is, it is desirable to develop domestic competency in the allocation of resources and management of risk, focusing on the particular facts and circumstances of the domestic economy. In the case of banks, even if domestic institutions were more competent (in allocating capital, managing risks, lowering transactions costs), foreign banks may have an advantage if domestic depositors or others that transact with the foreign institution *believe* that they are safer than domestic institutions.[6] They may be safer either because they are truly more sound, or simply because depositors (and others who transact with the bank) believe their richer governments stand behind them.[7] Those governments have the capacity and willingness to bail out their national banks, a belief that turned out to be largely true in the 2008 crisis but not in the Argentinean crisis.[8] Whether these beliefs are justified or not, if individuals hold these beliefs, then they will demand higher interest rates for deposits in domestic banks. In short, these beliefs put domestic financial institutions at a disadvantage.

Now, obviously, one would like to be able to test this idea empirically, and there is a set of data on which one can do so. The less-developed part of the United States economy includes the states of the old confederacy (Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, North Carolina, Tennessee, Texas, and Virginia). One can look at the relative development of those states in the period where the United States had only state banking and there was no interstate banking. That basically was the regime over the twenty years from 1850 to 1870. (There are other big changes over this period, most notably that have to do with the war, and that's why one wants to look at growth over a longer period.)

Between 1850 and 1870 the relative income, that is, the income in those southern states relative to total U.S. income, grew by 15 percentage points.[9] Growth was slightly higher in the poorer states of Alabama, Mississippi, and South Carolina than it was in the richer states of

Georgia, Virginia, and North Carolina. If one looks at the subsequent period of national banking, from 1870 to 1890, the catch-up is much, much slower: on average only 6 percent, and it is skewed toward the richer states: Virginia, Georgia, and North Carolina in particular. In fact, growth almost stops dead in its tracks in Alabama and Mississippi. Our hypothesis of what happened is simple: With local banking, local savings went to local entrepreneurs, and local financial institutions had incentives to develop the competency of judging among these. With national banking,[10] funds flowed out of the poorest states, which were afflicted with a low level of human capital and poor infrastructure.

Similar patterns seem at work today in developing countries. With financial market liberalization, foreign banks can attract funds away from domestic banks—or domestic banks will have to pay a higher interest rate to retain their depositor base. But, as we noted earlier, foreign banks have less information and less competency in judging among domestic small and medium-sized enterprises. Funds get diverted either back to the home country (about which they have superior information) or countries similar to or with close trade and financial ties to the home country. Any funds remaining within the developing country get allocated to seemingly more "secure" investments—loans to domestic monopolies and oligopolies and to the government.[11] To the extent that local banks are able to retain their depositor base by paying higher interest rates, the additional costs get passed on to local borrowers, putting local small and medium-sized enterprises at a disadvantage.

2. Capital Market Liberalization

Capital market liberalization (allowing the free mobility of capital in and out of a country[12]) has long been one of the most contentious areas of developmental policy. The IMF tried to change its charter, to allow it to force countries to liberalize their capital markets, in 1997—just as the East Asian crisis struck. The crisis was, in no small part, a result of the liberalization of capital markets in East Asia—which led to capital at first rushing in and then, when sentiment changed, rushing out (see Stiglitz 2002a; Furman and Stiglitz 1998). Interestingly, fifteen years later, in the aftermath of the global financial crisis of 2008, the IMF changed its position, recognizing that capital controls can play an important role in macroeconomic stability (IMF 2012b).

Ironically, capital market liberalization was supposed to bring both greater stability and growth. Capital would (in the standard theory that was part of the conventional wisdom and that underlay the Washington Consensus policies) flow into the country when it was needed—in a countercyclical way. And capital would flow from rich countries with a high capital to labor ratio to poorer countries with a low capital to labor ratio, enhancing growth in GDP in the latter and increasing global efficiency.[13] On both accounts, the standard theory failed: capital flows were largely pro-cyclical; overall, capital flowed out of developing countries; and, not surprisingly as a result, countries that liberalized their capital account had both more instability and lower growth.[14] (It was remarkable that these theories gained the ascendency that they did, given the overwhelming empirical evidence suggesting that capital flows were pro-cyclical.)

In the next chapter, we'll discuss briefly the impact of capital market liberalization on macro-stability. Here, we focus on the impact on the structure of investment—and thereby on learning.

What was so striking about East Asia was that with its very high savings rate, there was little need for capital inflows (unlike other parts of the developing world). What happened was that capital flowed in—and simultaneously flowed out. The countries typically borrowed at a much higher interest rate than they lent. They "invested" in U.S. government bonds, but borrowed from private (foreign) financial markets. In effect, capital market liberalization meant that they were outsourcing the allocation of capital to Western financial institutions. But, as our earlier analysis of banking suggested, these foreign financial institutions were less well informed about local conditions. In the case of several of the countries, money flowed particularly to real estate—helping create a real estate bubble. Thus, money flowed into sectors in which there was less learning.[15]

There were even more adverse effects on the structure of the economy (with respect to creating a learning economy) in the aftermath of the crises which regularly followed capital market liberalization (the East Asia crisis itself being but one example). Money flowed into these newly liberalized economies in a wave of irrational exuberance, but when attitudes changed, money flowed out. The sudden fall in the exchange rate created a currency crisis, which in many cases was translated into a financial and economic crisis. What happened next in Thailand illustrates the adverse effects. The IMF "rescue" package (designed to rescue the foreign lenders, not to protect the Thai economy) demanded high interest rates and large government cutbacks in spending. The result was a major economic downturn, with small and medium-sized domestic firms particularly hard hit and with major cutbacks in human investments (education and health). In effect, the attempt to arrest the fall in the currency—to enable firms that had borrowed in foreign exchange to be in a better position to pay back their debts—came at the expense of local entrepreneurs. In this case, the policy interventions were particularly unfortunate. The bail-out package had a barely noticeable effect on the exchange rate. Even if it had, much of the foreign-denominated lending was for real estate, and the real estate sector was already dead, with the breaking of the real estate bubble, and would not have been brought back to life even if the currency had been stabilized. But the high interest rates put a large fraction of smaller enterprises into bankruptcy—businesses that had never gambled by taking on foreign-denominated loans to finance activities, the returns to which would accrue in baht. And yet the damage to the growth prospects of the economy could be long lasting.

Capital market liberalization had a more direct effect on "learning," even in countries that did not experience a crisis. Because the allocation of capital was outsourced to Western financial institutions, there was less learning within the less-developed countries about the processes of capital allocation, risk management, etc.

Capital Outflows

Much of the discussion of recent capital market liberalization has focused on allowing capital inflows. But historically, restrictions on capital outflows have been even more important.

China restricted the outflow of capital, though it allowed an inflow of foreign direct investment. The effect of these capital market restrictions almost surely was to lower the cost of capital below what it otherwise would have been, facilitating domestic investment. If, as Arrow suggested, learning is related more to investment than to production, the result is more learning.

The general point (from the perspective of this book) is that if there are externalities associated with investment, when the money is invested abroad, those within the country don't benefit (at least as much) from these externalities.

3. Finance and Industrial Policy

One of the reasons that markets fail to allocate resources efficiently to "learning" is capital market constraints and imperfections. R & D is hard to collateralize,[16] and optimal learning entails expanding production beyond the point where price equals short-run marginal costs. Imperfections of information often lead, especially in developing countries, to credit and equity rationing.

Government regulation of the financial sector—and in some cases, direct government control over financial resources—can be effective instruments for directing resources to the creation of a learning economy. Even in the United States, usually viewed as a country in which the private sector dominates finance, at times government has played an active role in directing financial resources, e.g., through the Export-Import Bank and the Small Business Administration. Interestingly, a key instrument of industrial policy in East Asia was *access* to finance, often not even at subsidized rates (Stiglitz and Uy 1996). By steering finance to "learning" sectors, overall growth can be enhanced.

There are several other aspects of "learning" in the design of financial policy. In the following paragraphs, we discuss a few of the more important of these.

Financial Restraint[17]

Creating a vibrant financial sector focused, for instance, on lending to the industrial sector is not easy. Traditionally, banks focused on very short-term lending, trade credit, collateralized by the goods being bought and sold. At the core of successful lending is information—identifying who are good borrowers and then monitoring them to ensure that the funds go to where promised and that the borrower behaves in ways which enhance the likelihood of repayment. For reasons that are now well understood (see Greenwald and Stiglitz 2003), such information is best produced within institutions (banks or venture capital firms). As the U.S. precrisis attempt at developing a market-based system based on securitization shows so forcefully, capital markets are not a good substitute for such institutionally based financial flows.[18]

But within such institutions, there are strong agency problems (managers take actions to increase their own well-being, at the expense of other stakeholders, including shareholders), and the actions of financial institutions, especially large ones, give rise to significant externalities (when a bank fails, it has significant consequences for others). That is why governments all over the world subject banks to strong regulations. The question is: What kinds of regulations will be most effective in promoting long-term economic growth?

Much of the post-2008 crisis literature has focused on how to avoid excessive risk taking, to ensure macroeconomic stability. We focus here on a more basic question: how to ensure that finance flows to the "right" sectors and at terms that are pro-growth.

The East Asian countries, as we have already suggested, were remarkably successful in doing so. One instrument that they used was financial restraint: restricting both entry into banking and deposit rates *mildly*.[19] Financial restraint increased the franchise value of banks, and this encouraged them to undertake more prudential lending. The greater safety of the banking system more than offset the effect of the slightly lower interest rate, to ensure a strong flow of funds into the financial sector. (Governments in the region took other actions to encourage savings, so that even though real interest rates paid to savers were very low, savings rates were very high.)

There was enough competition within the banking system (accompanied perhaps by government pressure) that the lower deposit rates got translated into lower lending rates,[20] while incentives (and again, government pressure) ensured that much of the spread got reinvested in the banks, enabling them to expand lending in subsequent periods. The greater franchise value of the banks also induced them to act more prudently.

In short, government policy led to a more stable banking system, one which had the capacity to focus on more long-term lending. But government policy went further in encouraging lending to the learning sectors.

Access to Credit for Learning Sectors

As we have noted, because of the pervasiveness of externalities and other market failures, funds do not necessarily flow to sectors which might be most conducive to long-term growth, e.g., as a result of learning externalities. Regulatory policies can be used to shape the flow of funds, for instance, by putting restrictions on the maximum fraction of funds that can go into real estate or into consumer loans, or on the minimum fraction of funds that can go into, say, the industrial sectors. Tax and regulatory policies (such as those affecting derivatives and credit cards) can also encourage banks to devote both human and financial resources to wealth-creating lending activities, rather than to speculative activities or to predatory lending.

We noted earlier the dependence of the flow of funds on information, and government policies can help shape the flow of information, e.g., by insisting that banks open up branches outside the capital city.

Broad-based restrictions (as opposed to more narrowly focused policies that direct lending toward a particular firm) have a political economic advantage: The bank is not directed to put funds into a particular firm. It has to find, within the given category, the best set of borrowers, i.e., those most likely to repay. As we noted earlier, there is a trade-off. Such broad-based regulations may not succeed in targeting firms with the highest learning externalities, but the political economy advantages may more than offset the reduced effectiveness of targeting.

In some cases, however, relying on the private sector to allocate funds in ways that are most socially productive (especially when learning is important), using regulation to curb abuses and to partially correct the market failures arising from pervasive differences between private and social returns, may be less effective than government directly allocating funds itself. (The presumption that the private sector was a better allocator of scarce capital was clearly undermined by the events leading up to the 2008 financial crisis.) Thus, many developing countries have found development banks to be an important and effective instrument for managing the flow of funds to projects and enterprises which are viewed to have large externalities to others. While earlier development banks had been criticized for making politically connected loans with low social payoffs, in more recent years, countries around the world have found ways of managing these agency and political economy problems.[21]

There are some "half-way" houses, where the government provides some of the capital or absorbs some of the risk associated with lending to certain sectors (small businesses, R & D ventures). Government might, in particular, absorb some of the macroeconomic risks, forcing the bank to focus on commercial risks—the viability of the project, on the assumption that the government maintains the economy at or near full employment.[22] These mixed systems are designed to curb the political economy problems that sometimes arise with government programs: Funds are only provided if there is some private lender that is willing to bear *some* of the risk in return for some of the return.

Access to Credit for New Entrepreneurs and Small and Medium-Sized Enterprises

Emran and Stiglitz (2009) have stressed learning about who is a good entrepreneur. The problem, as we noted earlier (chapter 3), is that because of "poaching," the benefits of identifying who is a good entrepreneur may not be appropriated by the lender; those who establish themselves as good entrepreneurs will be recruited away by other banks (or firms). Even if the initial lender retains them, it will be forced to lend at a competitive interest rate—too low to recoup the losses incurred in lending to untried entrepreneurs, some of whom turn out to be incompetent.[23] There will be too little lending to new entrepreneurs. Government regulations can, however, mitigate the problem, e.g., by imposing restrictions on new entry into banking. But there is an economic cost of such regulations, and it may be preferable for governments to intervene more directly. That is one of the reasons that governments in many countries (including the United States) have government-funded programs to encourage lending to small and medium-sized enterprises. In some cases, these entail lending directly; in others, the absorption is part of the risk.

The learning externalities associated with small and medium-sized enterprises may be particularly significant, because of the high turnover of such firms. There is a higher likelihood that the personnel from such firms wind up in a relatively short time in some other firm, carrying with them the knowledge that they have acquired.

4. Concluding Comments: Restrictions on the Movement of Labor

This chapter has questioned the contention that the free and unfettered movement of capital—within and between countries—is desirable. We have shown that there is a role for government financial regulations, not only in promoting greater macro-stability but also in affecting the structure of the economy. Financial and capital market liberalization has often resulted in fewer resources going into sectors generating learning externalities and has therefore produced lower economic growth. There are likely to be stronger spillovers from lending by domestic financial institutions than from lending by foreign financial institutions. Restrictions on capital flows and on the provision of financial services by foreign firms are often effectively restrictions on the importation of foreign financial services, and just as we have argued that restrictions on the import of industrial goods may be desirable, there is an even more compelling case for restrictions on the import of foreign financial services.

Financial policy can also be an important tool for industrial policy. But even when the government is not directly using the financial sector to encourage some sector, or discourage another, broader financial sector policies can be used to help create a financial sector that is more focused on lending, and especially lending to activities that generate learning externalities. We have also suggested that while regulation of private financial institutions may be an effective tool for accomplishing these goals, government actions may need to go beyond this, creating government-run institutions (development banks) or programs that provide partial guarantees to private institutions.

Similar arguments, perhaps with even more force, though probably more controversially, apply to the free movement of labor across borders. Consider, for instance, the problem posed by the brain drain.[24] It is probably not a coincidence that the successful history of growth in Latin America seems to evaporate about the time when U.S. immigration laws changed, in 1965. In effect, those new policies enabled them to "import jobs from" (that is, get jobs in) the United States for their best and most talented people, who then don't disseminate the learning on a local basis, as they would have done if they had stayed in Latin America.

We should emphasize that we are not necessarily arguing for restrictions on out-migration, only noting that there are growth consequences—externalities—that should not be ignored. There are policies that could mitigate some of the costs, e.g., international agreements that would require those emigrating from a country to repay investments in human capital. But while these policies partially address the fiscal consequences of migration, they do not adequately address the more fundamental problem of learning externalities.

At the same time, there are some instances when returning migrants bring back with them not only money (which has often been emphasized in the recent literature on migration and development) but also knowledge—sometimes called "cultural remittances." Migration *can*, if properly structured, facilitate cross-border movements of knowledge and learning. It can thereby help create a learning economy.

CHAPTER ELEVEN

Macroeconomic and Investment Policies for a Learning Society

CHAPTER 4 explained the key reasons that macroeconomic volatility is bad for a learning economy, and accordingly, why policies which exposed countries to more volatility can have adverse effects on learning—quite apart from whatever other positive or negative effects are associated with such policies.

This chapter looks at four broad issues regarding how government macroeconomic and investment policies can help create a learning economy. The first looks at financial and capital market liberalization. The last chapter explained why such policies may not be desirable, as a result of their impact on the *structure* of the economy—they change the structure in ways that slow learning. Here, we use an information-theoretic analysis to help explain why such liberalization may systemically be associated with more macro-instability. The second looks at exchange rate policy. With the advent of increasing international restrictions on the use of trade and industrial policy, adjustments in exchange rate have become an increasingly important tool. Moreover whether and how the government manages the exchange rate—both the level of the rate and its volatility—affects the structure of the economy and the extent to which there is learning, and thus growth. The third looks at foreign direct investment (FDI), asking how and when it can be used to promote a learning economy. (The previous chapter and section 1 of this chapter focus on *financial* flows, not FDI.) The fourth examines the role of government investment.

1. Financial Policy and Macro-Stability

It has been widely noted that capital market liberalization is often associated with greater macroeconomic instability. To advocates of capital market liberalization, especially those wedded to the standard neoclassical model assuming perfect markets (including the absence of any information asymmetries), this seemed a puzzle. Of course, for anyone observing actual patterns of capital flows, it would have been more of a mystery had it turned out that capital market liberalization was stabilizing, for such flows tend to be pro-cyclical.

In the discussion that follows, we suggest one of the reasons that capital market liberalization is often associated with instability.

Imagine that a foreigner invests his money in a country, alongside local investors. The local investors, let's say in the banking system, know a lot more about what's going on than foreigners do. They have a better sense of what difficulties (either economic or political) can be easily resolved and which cannot be. When a firm fails, the local investors can better ascertain whether the failure was due to idiosyncratic events confronting the firm, or whether it was due to systemic effects likely to affect other firms. So when there is trouble within a country, local investors will know to take their money out of the country before foreigners do. But foreigners know that they are in this disadvantageous position; they know that there are these information asymmetries and that they observe things only with error. Accordingly, they set a lower threshold at which they take their money out—even slightly adverse signals may lead to large capital outflows by foreigners.

In practice, that has three implications. The first is that countries that restrict capital movements will have more stability (see, e.g., Stiglitz and Ocampo [2008] and the articles cited there).

Second, crises will be *disproportionately* generated by overseas investors because they are the uninformed investors who react often even to small signals.[1, 2]

But the third and most surprising implication of this is that the countries that are successful at "exporting" financial services (that is, to selling their financial services to outsiders) ought to be small countries, because in small countries there are fewer locals who are at an informational advantage over foreigners. These issues of instability and redistribution between the locals and the outsiders will be less severe. This perhaps explains why it is the small countries like Switzerland, Luxembourg, and Iceland that turn out to be (disproportionately) the financial centers. (However, as several of these countries have learned at great expense to themselves, there are real costs, including to macro-stability, from having an outsized financial sector.)

Financial Market Liberalization

Financial market liberalization, like capital market liberalization, was supposed to enhance macroeconomic stability. The argument was that strong international banks would be able and willing to continue the flow of funds to, say, a developing country whose own financial institutions were going through a period of weakness.[3] But there is evidence that it has not done so (Rashid 2012). One of the reasons is that just noted: foreign banks are at an informational disadvantage and so are more prone to pull money out of a country in response to adverse signals.

But another is that foreign banks can be a source of economic volatility, especially if they come to take a large share of the domestic financial sector. Because foreign banks are less informed about foreign borrowers, they view such lending as riskier. When banks face an adverse balance sheet shock (e.g., in the event of an economic crisis at home), they become more risk averse. They respond by pulling back most on the riskiest lending—including lending abroad (Greenwald and Stiglitz 1993, 2003).[4] These responses become even stronger under home government pressure, when the government has put in funds to rescue the bank (or potentially might do so). The justification for such bailouts is that they are necessary to maintain the flow of credit *within the country*, but citizens will then expect that the flow of credit in fact go to their own country, not abroad. Hence, there will be demands that banks allocate scarce lending resources to home country borrowers. There is some evidence that this occurred during the 2008 global financial crisis.

The consequence, of course, is that financial market liberalization may lead to the transmission of problems in one country to another. It can be an important way for macroeconomic instability to move across borders. In fact, there is considerable evidence that much of the instability in developing countries during the era of global financial market liberalization came from outside the boundaries of the countries: Volatility in capital flows and global risk premiums have become an important source of macroeconomic instability.

2. Exchange Rate Policy

Exchange rate policy is concerned about both the level of the exchange rate and its volatility.

The exchange rate affects the competitiveness of the economy—the ability of exporters to export and of import-competing firms to compete with imports. The consequences of an appreciation of the currency (say, as a result of the inflow of capital or foreign aid) can be severe: If the exchange rate increases by, say, 25 percent, there is no way that (in the short run) productivity can compensate or for there to be offsetting adjustments of wages and the prices of other inputs. Moreover, there are important hysteresis effects (just as there are with large movements in the interest rate): A firm that dies because it can't compete as a result of a high exchange rate is not brought back to life when the exchange rate subsequently falls. (Capital market imperfections imply that especially small and medium-sized firms will be unable to obtain the capital required to tide them over.)

Exchange rate volatility is closely associated with macroeconomic instability; it is both cause and consequence. In a world of flexible exchange rates, changes in beliefs about the future value of the exchange rate can induce large financial flows and lead to large changes in the exchange rate today, with large macroeconomic consequences, e.g., for trade.

By the same token, it is expensive for firms (especially small and medium-sized domestic firms) to manage exchange rate volatility, especially in emerging markets and in the least-developed countries. In many of these countries, there may be no markets in which firms can hedge their exchange rate risks. But even if there are markets for hedging, there is a cost to doing so.

All countries have an exchange rate policy—whether they like it or not. Even a decision not to intervene in the exchange rate directly is a *policy*—one which subjects the country to high levels of volatility and lets the exchange rate be determined by others. The result is that countries like South Africa that have adopted such policies have had among the most volatile exchange rates.

Some have argued that one should let exchange rates be determined by market forces. But, again, like it or not, government policies are at the center. When a country raises or lowers the interest rate, it affects the exchange rate. When a country changes a regulation affecting flows of money into or out of the country, it affects the exchange rate. Since most countries have a broad array of regulations affecting inward and outward investment, there is, in a sense, no "free market" exchange rate. Through these regulations, interest rates, and direct interventions, governments set the exchange rate, intentionally or not.

Most importantly, when a large country like the United States changes its policies, it affects the exchange rates of almost all others. Thus, when the United States first adopted quantitative easing, it led to an appreciation of currencies around the world, and when it began tapering, that too had global consequences. The question facing any country is not whether they should let the exchange rate be determined by market forces, but whether they should let their exchange rate be determined by the Federal Reserve Board.

And, exchange rate policy affects the industrial structure. Some sectors, some technologies, and some kinds of firms are discouraged relative to others. A decision not to actively manage the exchange rate will result in a more volatile exchange rate, more macroeconomic instability, and a smaller traded goods sector more dominated by large firms than would otherwise be the case. In the context of Africa, the decision of many resource-rich countries to allow their exchange rate to appreciate has contributed to deindustrialization and even the weakening of the agriculture sector.

Exchange rate policies are both a tool of industrial policy itself (as we have explained in earlier chapters) and a *complement* to other industrial policies. Policies designed to encourage the manufacturing sector won't work if the exchange rate is not competitive or is highly volatile. Some in South Africa observed that even though the exchange rate had fallen markedly in early 2014, the manufacturing sector had not expanded. But that was totally understandable; given the history of volatility, no firm could count on the exchange rate to remain competitive, and it would be foolish for any firm to enter based solely on the *current* exchange rate.

There are several implications of this analysis. First, governments need to adopt policies that make exchange rates less volatile, e.g., capital controls (or more generally, they have to adopt a portfolio of tools for capital account management).[5]

Second, governments need to keep exchange rates "low" so as to make domestic firms more competitive—to expand exports and import-competing sectors. This is especially true because low exchange rates help export sectors like manufacturing, which have higher learning elasticities and generate more learning externalities. (There may be further macroeconomic—and learning—benefits if the lower exchange rate allows the country to operate closer to full employment.[6])

But a concern about industrial policy means governments need to be attentive to *how* they intervene to stabilize and lower the exchange rate. If to prevent a large decline in the exchange rate they increase interest rates (as was the wont of the IMF), while they may thereby save large numbers of enterprises that have taken on foreign denominated debts, at the same time they may kill other enterprises that were more prudent and took on only domestic debt. The effects may be particularly adverse for small and medium-sized enterprises (which typically do not take on foreign debt, because they do not have access to international markets)—as was evident in the East Asia crisis (Furman and Stiglitz 1998). Again, if these small and medium-sized enterprises play an important role in creating and disseminating knowledge and learning, such policies can be particularly counterproductive. In any case, governments should be aware that whether and how governments intervene in managing the exchange rate has implications for economic structure—and for learning.

There are alternative ways of stabilizing the exchange rates and, even more so, keeping exchange rates low, which may be less costly. Low interest rates benefit domestic firms on two accounts, both directly and indirectly, through effects on exchange rates. Direct intervention, with the consequent buildup of reserves, may be a particularly effective way of smoothing the exchange rate, at the same time avoiding volatility in interest rates.

Some have suggested that it is impossible to push the *real* exchange rate down for more than a short period of time. But such arguments are based on a confusion. It is impossible to keep exchange rates above the "market" level through direct intervention, because to do so requires selling dollars (or other hard currency), and countries only have limited amounts of these in their reserves. But to push the exchange rate down requires selling one's own currency and buying dollars (or other hard currencies), and this countries can easily do.

There are many instruments available for affecting the *level* and volatility of the exchange rate.[7] As we have noted, any regulation that affects the flow of money out of or into the country affects the exchange rate. Thus, making it easier for foreign companies to invest in the country leads to the appreciation of the currency; making it more difficult leads to the depreciation of the currency. In assessing foreign direct investment policy, one has to weigh the benefits of access to markets or technology or training with the costs to the rest of the economy from the exchange rate appreciation (including the adverse effects on learning). Capital market liberalization similarly can lead to an influx of capital and an appreciation of the currency. By the same token, loosening restrictions on citizens of the country investing their money abroad lowers the exchange rate, making it more difficult or less attractive for short-term money to flow easily in and out of the country, thus contributing to volatility.

A lower exchange rate represents a broad-based mechanism for industrial policy—firms themselves decide whether they can compete at that lower exchange rate. The government has identified broadly that the export sector has more learning externalities and therefore that sector should be encouraged relative to others, but it doesn't have to identify precisely which subsectors or firms should be encouraged. The market does that.

This has an advantage and a disadvantage, which we have briefly noted in earlier chapters. More finely tuned targeting may increase the overall (dynamic) efficiency of the economy; after all, each firm or sector takes no account of the extent of the benefits that accrue to others. A more targeted approach can offset the externality associated with research or learning in each sector. On the other hand, government attempts at fine tuning may encounter more severe "political economy" problems.

There are two questions about the use of these instruments. First, what really matters is the real exchange rate. The question is: Can government affect, at least more than just briefly, the real exchange rate? Here, the critical question is the extent and speed of "pass through."

For very open economies, importing and exporting a large fraction of their goods, lowering the nominal exchange rate leads to increases in nominal prices, which can undo the benefits, unless, say, monetary authorities take actions to dampen the potential inflation, but such actions themselves have costs (e.g., higher unemployment). It is clear, however, that many countries have managed to lower their *real* exchange rate for an extended period of time and have done so at the same time that they have promoted growth.

Second, what are the costs of each of the interventions, and do the benefits exceed the costs? Some worry that the cost of preventing inflation from direct intervention is too high. The East Asian countries have managed to intervene in the exchange rate over long periods of time without facing either high inflation or high costs of avoiding inflation. But, at least in China, there is another growing concern: to keep the value of their currency low, they have bought dollars, which yield a low return. Worse, dollars are depreciating relative to the RMB, implying that they are experiencing a (paper) capital loss.

But the cost of reserves has often been exaggerated. The reserves are created by selling domestic currency (which costs nothing to produce) for dollars. If there were no exchange rate management, there would be no reserves. The main real cost arises from the differences in the exchange rates between the time the dollars are bought and the time they are sold. And there are policies that can increase the return to reserves. For example, if the reserves are going to be maintained for a long time, hold long-term rather than short-term debt, and rather than buying government bonds, buy a diversified portfolio of stocks.[8]

Industrial policies can intervene in relative prices in ways which avoid these costs (and which can in fact be more targeted than lowering the [real] exchange rate), e.g., by sectoral subsidies (including subsidized interest rates) or infant-industry protection. But, as we noted earlier, international trade agreements restrict the use of industrial policies. The only instrument left may be the exchange rate.

In any case, these costs are typically small compared with the benefits. Indeed, it can be shown that benefits are so large that it would pay countries to accumulate reserves *even if they did not subsequently use them*. Essentially, this seems like throwing money away, but as countries grow richer and their populations age, it is likely that the reserves they have built up—and which facilitated their transition to a modern learning economy—will prove invaluable.

3. Investment Policies

In some (but not all) of the successful countries, foreign direct investment (FDI) has played an important role. For some countries with limited access to finance, FDI can be an important source of funds. But even in those countries with high savings rates, champions of FDI extol its virtue in terms of the transfer of knowledge. But this doesn't happen automatically; the learning spillovers are more important for some forms of FDI than others; and there are ways of transferring knowledge other than by FDI. While FDI may be one way of acquiring knowledge, there may be other ways of doing so, which simultaneously induce more learning. Whether that is the case may depend on the admissible rules governing FDI (which we will discuss shortly).

It is worth noting that as we look around the world at countries that have been most successful in development, some have succeeded with little FDI (Korea and Japan), while in others, FDI has played an important role. For the most part, though, the countries that have relied on FDI have not succeeded in creating global enterprises of the likes of Samsung, Toyota, and Sony.

Foreign firms have both advantages and disadvantages in promoting learning. On the one hand, knowledge often moves more freely within the boundaries of the firm than across firm boundaries. Within-firm knowledge may thus be the most effective way of moving knowledge across national boundaries.

On the other hand, firms actively work to limit the knowledge so transferred from seeping out beyond the confines of the firm. In spite of such efforts, though, knowledge is transferred—and this is in fact one of the main justifications for encouraging FDI.

But firms may feel that keeping knowledge within the confines of the firm may be more difficult outside the home country, and thus they may restrict the flow of knowledge across borders, consigning, say, developing countries to using less advanced technologies. In this case, FDI may not be the best way for developing countries to close the knowledge gap. (Western companies and governments sometimes claim that they would more readily use more advanced technologies in developing countries if these countries more effectively enforced intellectual property rights. But there is little convincing evidence to support this hypothesis. The next chapter discusses intellectual property rights more broadly.)

In many, perhaps most, cases, entrepreneurial spillovers may be larger in the case of domestic enterprises than foreign, since domestic firms are likely to be more firmly embedded within the local community. Government policy should, accordingly, provide some preference for domestic firms relative to foreign firms, except when there are strong learning benefits that are specifically related to foreign firms, e.g., because the foreign firm brings knowledge that is not locally available and can't be otherwise acquired.

One question that governments attempting to create a learning economy through FDI need to ask is: What kind of FDI is likely to be most beneficial? The theory of localized technological change (Atkinson and Stiglitz 1969), discussed in chapter 3, explains that the spillovers from learning associated with one technology are more likely to be greater for "nearby" technologies. What matters is both the *relevance* of the knowledge associated with one technology for the improvement of another and the *capacity* of those employing one technology to learn from another. Accordingly, spillovers may well be stronger across sectors for similar technologies than within the sector for markedly different technologies. Thus, just-in-time inventory practices have benefits for many sectors in which inventories play an important role.

Much of the knowledge which is embedded in, say, mining technologies is of limited relevance to most other sectors of the economy. Thus, the learning benefits of FDI associated with resource extraction are likely to be much more limited than those associated with, say, manufacturing, and this may help explain why so many resource-dependent economies remain "dual" economies, with few spillovers from the natural resource sector to the rest of the economy. If this is so, it means that FDI in this area—one which has dominated in Africa—is of much less benefit than FDI in other areas. (But as chapter 9 emphasized, there are still ways in which industrial policies can play an important role in natural resource economies. The limited spillovers are, in part, the result of the absence of adequate industrial policies.)

By the same token, the benefits of export-processing sectors are often disappointing (beyond the direct job creation and limited tax revenues that they generate). Weak links to the rest of the economy mean that spillovers are limited, and the low-technology, labor-intensive processes in which they specialize would, in any case, give rise to few spillovers. At the same time, government-sponsored industrial and research parks, promoting clusters of related activities, can promote growth by generating and capturing learning spillovers.

While it may be easiest to learn about adjacent technologies, the benefits of such learning may be more limited than that associated with making larger steps (sometimes referred to as leapfrogging). There is, then, a complicated optimization problem: Both the costs and benefits increase the larger the step. Moreover, one wants to move toward technologies from which one can learn the best going forward, and that may not always be easy to assess from one's current vantage point. Korea and Japan's industrial development was characterized by strategies that involved moving some distance from the technologies that they were then employing.

Government Subsidies for and Regulation of FDI to Promote Learning

Government subsidies for FDI have typically been justified in terms of the government revenue and employment generated. But our analysis suggests another rationale: learning. But if this is so, then subsidies should be larger for those sectors and technologies which are likely to have large spillovers and for firms that are willing to engage in practices that enhance the likelihood of such learning.

Government policies can also affect the extent of spillovers. Compulsory employment and training programs and domestic procurement requirements (programs that compel firms to source locally) are more likely to lead to learning spillovers. The success of Malaysia's FDI was partially attributable to these requirements. The benefits of learning can more than offset the social costs of the distortion.

It is worth noting that trade and investment agreements have typically circumscribed the use of these kinds of interventions and have tended to insist on foreign firms being given equal—or in some cases preferable—treatment with domestic firms.[9] Like other aspects of Washington Consensus and neoliberal policies, these provisions need to be reexamined through the prism of learning.

Outward-Bound Investment

China and other emerging markets are increasingly engaged in buying foreign firms. This can be seen as an alternative, and sometimes more effective, way of acquiring knowledge. Ownership of a firm gives them the right to transfer the technology. The new owners may seek to acquire the technology and transfer production to their country, where the subsequent learning will occur. The result is that the country that had been the technological leader in a particular arena will lose its dominant position. More broadly, from the perspective of societal learning, the locus of such learning will change too: More will occur in the country of the acquiring firm.

There is much to be said for China's learning strategy: First it developed the ability to learn, through investments in education and foreign direct investment in China. As foreign governments and firms saw China and Chinese firms as more of a competitive threat, FDI became a less effective way of learning (at the margin) compared to buying foreign firms—which became increasingly feasible with China's mounting resources. At this point, for the most part, the "rules of the game" worked in favor of China—there were relatively few restrictions on the acquisition of firms, even when the intent and consequence was clear, the acquisition of technology, and the shift in the locus of production. (An important exception was U.S. restrictions on the acquisition of high-technology industries. But here, U.S. restrictions on high-technology exports provided China with an even greater impetus to develop its own capacity, acquiring technology from other countries not imposing such restrictions.)

4. Government Investment and Expenditure Policy

We have argued that governments cannot avoid questions of industrial policy, for they have to make decisions about the direction of public investment, say, in education and infrastructure. These decisions have to be based on beliefs about the future directions of the economy, which are in turn affected by these public decisions. But the policies with which we are concerned go well beyond this. For government can use public expenditure policies to partially compensate for deficiencies in market allocations. It can provide infrastructure and public goods which are complements to learning sectors. It can design government procurement policies which facilitate learning—demanding that the products bought by the government are produced at home, especially when this production is associated with high levels of learning and learning spillovers. (Many signatories to the WTO have not signed on to the government procurement agreement; in fact, only forty-three have. Countries should be forewarned that this agreement may limit their ability to implement these learning strategies.)

This chapter has explained the important role of government in establishing an environment that is conducive to creating a learning economy, through public investments that support private-sector investments in learning sectors or by creating a stable macroeconomy.

Virtually every policy of the government can be reexamined through the prism of learning, but doing so would take us beyond the confines of this short book. We should, however, mention four in particular.

Education

The first is education. Nothing is more important than educating young people for creativity, and an education system that is focused on learning how to learn and on lifetime learning. Much attention is focused on learning in our educational institutions during the early years of an individual's life, but far more important is the education that occurs elsewhere, including on the job. Recognizing the complementarity between these different forms of education and designing formal education systems and on-the-job training programs that are complementary to each other and which, together, maximize overall learning are essential to creating a learning economy.

Only a small part of learning occurs in formal education, but government has to make sure that it sets the appropriate pre-conditions for learning later in life. The relationship between the two is changing with the increasing pace of innovation, changes in the labor market, and changes in technology. For instance, in the "old" model, students were stuffed with knowledge that teachers hoped would be useful in later life, partly as a building block for future learning.

Because of the fast pace of change, knowledge acquired in school may be of limited relevance 20, or even 10 years later. And today, through the Internet, individuals have access to far more knowledge than any teacher could hope to stuff into even the most receptive student. Education needs to focus learning on how to access, appraise, and make sense of this wealth of information.

In the "old model" good firms would provide on-the-job training. It paid for them to do so because employee turnover was typically low.

Now turnover is high, and firms will be able to appropriate for themselves only a small part of any education and training they provide. The burden of learning will according have to shift to the individual himself and to the government.

Fortunately, there are developing new technologies (such as Massive On-line Open Courses, MOOCS) making high-quality, free education available over the Internet. How this plays out in the future remains to be seen, but these new technologies and institutional arrangements may enable individuals to tailor learning programs more to their interests, capabilities, and perceptions of the changing marketplace.

This learning perspective has had a particularly profound impact on thinking about education in developing countries. The "old model" argued that both efficiency and equity called for prioritizing the provision of primary education for all.

But as we have emphasized, it is now recognized that what separates developed from developing countries is as much a gap in knowledge as anything else. The education system has to be designed to close that gap. That means that there has to be a sufficiently large number of individuals with secondary and tertiary education to absorb knowledge from the advanced countries.

Even primary education systems have to be redesigned from the learning perspective. Learning basic reading, writing, and arithmetic skills is necessary but not sufficient. Learning the history of the kings of England (part of a traditional colonial/ex-colonial education) is not as important as learning the benefits and risks associated with different fertilizers. Children have to learn skills that are relevant to the contexts in which they live. With the vast majority in rural Africa likely to continue to live in the rural sector, education has to be directed at improving lives in the rural sector—not just qualifying individuals for urban jobs that may not exist.[10]

We have a little more to say about education in chapter 13.

Social Protection

The second policy area is social protection. We don't typically think of systems of social protection as part of our "learning society" policy framework, but they are—for learning is risky. We argued here and in chapter 4 that risk has an adverse effect on creating a learning society. What matters, though, is not just the level of risk but how risk is managed. The adverse effects of "trying and failing" are lower in a society with a good system of social protection. In short, there can be more risk taking in societies with better systems of social protection.[11]

Legal Frameworks

There are other aspects of a country's legal and economic framework that can affect risk taking as well. Bankruptcy laws give an individual a fresh start, and thus good (and especially debtor friendly) bankruptcy laws can encourage risk taking. Such bankruptcy laws also may encourage better lending practices. Many believe that it is not an accident that the worst lending in America's mortgage market occurred as the United States "reformed" its bankruptcy laws to make the discharge of debt more difficult (Stiglitz 2010b). And as we have previously noted, a badly designed bankruptcy law, combined with deficient regulation of for-profit predatory educational institutions and the high cost of public education, discourages investment in tertiary education.

Innovation System

An important aspect of a country's economic and legal framework that is crucial in determining whether the country becomes a learning society is a country's "innovation system," including the set of laws and institutions that support explicit investments not only in education but also in research. Most important are investments in basic research by the government, which do much to define the opportunities available for others to make advances. How those advances are incentivized and financed, and how knowledge is shared, has much to do with creating a learning economy. A central feature of any country's innovation system is the country's intellectual property regime, the subject of the next chapter.

CHAPTER TWELVE

Intellectual Property

INTELLECTUAL PROPERTY provides a final application of the general ideas we have attempted to develop in this book. Intellectual property rights (IPR) are supposed to provide incentives to encourage innovation. As such, a discussion of IPR fits squarely into any analysis of policies centered on creating a learning society. Our concern here is that the provisions of the intellectual property regime that has become dominant around the world (reflected in the TRIPS provisions of the Uruguay Round global trade agreement) do not maximize learning. This was evident even at the moment of their creation. They were explicitly not driven by a broad focus on enhancing societal innovation or even well-being (though the rhetoric advanced in their support sometimes suggested otherwise). They did not maximize learning and the pace of progress for the United States and other advanced industrial countries; they were even worse for developing countries. Rather, they were designed to maximize rents of the entertainment and pharmaceutical industries.

In this chapter, we attempt to explain some of the reasons that these standard IPR provisions do not maximize learning and what might be done about it. And while we shall have much to say about the consequences for developing countries, there is dissatisfaction even within the United States. Many in the software industry, for instance, have been agitating for changes. As we note at the end of the chapter, there have in fact been some dramatic changes to the U.S. intellectual property regime in recent years, though mostly through court rulings, rather than through legislation.

1. IPR and the Relationship Between Social and Private Returns

Given our emphasis in this book on learning externalities, it might be thought that we would support the conventional wisdom, that the stronger the intellectual property rights, the better. For at least with strong intellectual property rights, it would seem that more of the returns to the innovation would be captured by the innovator. This would reduce the gap between social and private returns.

The problem is that the disparity between private and social returns is multifaceted, and, as the discussion of part 1 hinted, stronger IPR may not lead to more innovation and, more broadly, may lower social welfare. Later sections of this chapter will explain that there are other ways of financing and incentivizing innovation that may lead to a faster pace of innovation or higher levels of social welfare. The problems are exacerbated with poorly designed IPR regimes; some of the problems with current IPR regimes could be ameliorated through reforms in the IPR regime discussed in the next section.

We can organize the distortions introduced into the economy system by IPR into two categories: IPR gives rise to a static inefficiency, and (especially inappropriately designed) IPR regimes may even impede the pace of innovation.

Static Inefficiencies

As we noted in chapter 6, knowledge is a public good, so any restriction on the use or dissemination of knowledge introduces an inefficiency. Moreover, IPR gives the owner of the knowledge (patent) exclusive rights over the use of that knowledge. It thus confers monopoly power over knowledge, which can give rise to monopoly power in production, which, in turn, introduces a distortion in production.

There are some circumstances in which these static costs can be especially high. For instance, monopoly drug prices make lifesaving medicines unaffordable to the poor in countries without public health insurance, resulting in large numbers of unnecessary deaths. And in developing countries with limited budgets, when governments pay monopoly prices to the drug companies, there is less left over for other health needs or to pursue broader developmental objectives.

Typically, advocates of IPR argue that a well-designed IPR regime balances these well-recognized static costs against the dynamic benefits of faster innovation. But, as we now explain, IPR may not even result in faster innovation.

Dynamic Inefficiencies

In chapters 5 and 6, we explained why a monopolist may have less incentive to engage in innovation than is socially optimal, simply because production is lower. Here, we explain further why a strong intellectual property regime may not lead to rising living standards. First, it may not lead to more innovation. Second, the innovation that does occur may not be well directed.

DOES STRONGER IPR LEAD TO MORE INNOVATION?[1] Earlier, we explained why Schumpeter's argument that Schumpeterian competition would provide a spur for innovation was wrong: The incumbent can get sufficiently far ahead of potential competitors that they would be discouraged from entering the fray; the monopolist could then simply rest on its laurels. Because stronger IPR can give rise to stronger monopoly power, these problems can be exacerbated by strong IPR regimes, which result in a less competitive marketplace.[2]

We have seen too how a monopolist may create entry barriers, engaging in a variety of entry-deterring practices that stifle innovation—as Microsoft so ably demonstrated. Its control of the PC operating system meant that it could (and in fact did) foreclose, or at least reduce, competition in the provision of applications. It took especially strong actions against innovations that were a threat to its dominance.

We also noted that, because monopoly restricts production, the level of innovation will be less than the socially optimal level (though possibly more than with competition, where the level of production of each producer is even smaller).

A major theme of this book is that there are always knowledge spillovers, but market participants won't take these societal benefits into account in deciding on the level of investment in innovation. This is even true under Schumpeterian competition, where there is a succession of monopolists, even when the innovation in one sector is of no benefit to those in other sectors. In this situation, the innovation that occurs at time *t* sets the base from which successor innovations take off, but each innovator fails to take into account the benefits it confers to the monopolists that follow.[3]

There are still other reasons that stronger IPR regimes may not lead to a faster rate of innovation, a number of which we mention briefly in this chapter.

Follow-On Innovation The most important is that knowledge is the most important input into the production of knowledge, into learning. Every innovation is based on prior innovations. Thus, to *maximize learning*, one must strike a balance between the benefits to the producer and the potential benefits to subsequent (follow-on) users.[4] Those in the drug industry, for instance, while they want to preclude others from using their innovations in follow-on research, seek to be sure to have the ability to use the innovations of others in theirs. These concerns were part of the original "deal" in creating our intellectual property regime: Those who seek protection of patents are required to disclose enough information to have the innovation replicated by others. Others could use that knowledge in their own learning, even if they could not use the product itself in their own developments. In practice, however, disclosure is typically inadequate, and some who have sought intellectual property protections have simultaneously fought hard against disclosure. (Microsoft is the most notable case. They have resisted disclosure of source code and, when required to do so, complied only after heavy fines have been imposed.)

Enclosing the Knowledge Commons Every idea builds on others. Patents are supposed to protect only new knowledge. But the boundaries of knowledge are not precise (in contrast to land, where we can demarcate boundaries precisely), and, inevitably, to some extent patents may extend to preexisting knowledge. To the extent that that is the case, they provide a return to the patent holder; but there is a *negative* societal return, since the patent will discourage innovations that might have made use of this otherwise freely available knowledge. Boyle has referred to such patents as "enclosing the commons" (Boyle 2003).

Patents of traditional knowledge, to be discussed further, are the most dramatic example of a patent extended to preexisting knowledge.

Of course, patents shouldn't be granted for preexisting knowledge, but it is often difficult to ascertain the boundaries of existing knowledge. The patented knowledge may, of course, not be known to the patent examiner, and, especially in the case of traditional knowledge, it may not have been published. It may not be easy to publish knowledge that is so widely known as to be considered "common knowledge."

Encouraging Secrecy The basic model of science—a model that has proved enormously productive—entails openness and the sharing of knowledge. But intellectual property not only interferes with the dissemination and transmission of knowledge, but it also encourages secrecy, which impedes learning. Indeed, the extension of IPR to universities, under the Bayh-Dole Act, has encouraged a culture which is antithetical to the openness that has traditionally characterized such institutions.[5]

Litigation Risk and Ambiguous Boundaries These problems are exacerbated by the absence of well-defined boundaries for knowledge, which we have already noted. Did or should George Baldwin Selden's original patent for a four-wheeled, self-propelled vehicle include all such vehicles or only the one he sketched out?[6] And because all knowledge is based on prior knowledge, one faces a difficult decision: When is a new idea really new, rather than a minor wrinkle on an old idea, or a different expression or representation of an old idea? Patent laws have addressed this question through standards of "novelty" and "obviousness." But there is inevitably ambiguity, and where the standard is drawn raises all the questions we have been discussing, balancing out the impediments to further innovation with the (alleged) benefits of induced innovation. There is thus always a critical issue in defining the scope of the patent.[7]

What is clear is that the patent offices have found it difficult to draw the line in ways which many, if not most, in the scientific community view as balancing the concerns appropriately. Too much weight is given to the current producer, too little weight to future users.

The difficulties posed by defining boundaries give rise to litigation. Lawyers have recognized the problems posed by these ambiguities and have had a field day. The litigation and the uncertainty surrounding litigation themselves become impediments to innovation.

The problem is compounded by the underfunding of the patent office, which limits the amount of time that patent examiners have to examine any patent, leading to the issuance of too many patents. Particularly before the eBay decision described later in this chapter, those who received these "weak" patents (patents that were unlikely to survive a challenge) were nonetheless able to extract considerable rents out of other innovators, with adverse effects on innovation (see Farrell and Shapiro 2008).

Patent Thickets[8] Today, most products are sufficiently complex that their production may require using many separate items of knowledge. If each piece of knowledge is protected by a patent, this engenders a complex bargaining problem. Unless the owners of these separate pieces of IPR can agree, the product cannot be produced.

The problems are even worse. Anybody engaged in writing a software program, for instance, even doing so with complete originality, faces the risk that in doing so she may have trespassed on one of the hundreds of thousands of related software patents or that she may have come close enough to trespassing to make herself libel to litigation. No one can keep up with the myriad of patents being issued—if one did, it would be difficult to have time to engage in research. In this sense, the patent system itself has become a roadblock to innovation.

There is a long history of patent thickets serving as obstacles to innovation. Early in the twentieth century, the development of the airplane was impeded by conflicting patent claims by the Wright Brothers and Glenn H. Curtis. It was only with World War I, when the government forced the creation of a patent pool, that progress occurred in the United States.[9, 10]

Holdups and Patent Trolls More recently a whole industry has developed—firms (called patent trolls) buy up patents, waiting until someone successfully produces a product that *might* have infringed on their patent. When they find such a product, they sue, in effect holding up the producer for ransom. To the extent that they can get more for themselves, there is less left over for the real innovators.

The holdup problem has been exacerbated by provisions in our patent system that allow the owner of a patent—even a patent whose legitimacy is under question—to exclude anyone else from using that knowledge.[11] The BlackBerry suit shows what can happen. The company was held for ransom by the owner of patent(s) that were under dispute. BlackBerry proposed a settlement, widely viewed to be fair (or even overgenerous), suggesting that if it turned out that the patents were not valid, a fraction of the payments be returned to them. The offer was refused. BlackBerry had no choice: It had to either accept the terms or shut down.

In the United States, a pathbreaking decision of the Supreme Court in a case called *eBay Inc. v. MercExchange, L.L.C.* may have profound implications for such suits.[12] In the past, patents have typically been enforced through injunctions—others cannot trespass on a patent without the permission of the patent holder, who can extract as much "rent" as he wishes. This is in contrast with many other areas of the law, where there is compensation for violating someone's rights or property. The Supreme Court itself has raised questions about the consequences of what might be termed excessive enforcement, when patent holders take actions that in effect exclude from the market those who *might* infringe upon the patent. In *eBay*, the Court ruled that a permanent injunction (against infringement) would only be granted if a four-part test was satisfied:[13]

> A [patent] plaintiff must demonstrate: (1) that it has suffered an irreparable injury; (2) that remedies at law, such as monetary damages, are inadequate to compensate for that injury; (3) that, considering the balance of hardships between the plaintiff and defendant, a remedy in equity is warranted; and (4) that the public interest would not be disserved by a permanent injunction.

An extreme version of exclusion to protect intellectual property is still part of America's trade laws, where a firm that the International Trade Commission finds has violated an American's intellectual property rights can have the infringing products excluded from importation into the United States.[14]

There are other reasons that stronger IPR may not lead to a faster rate of increase in standards of living. Even if stronger IPR increased investment in innovation, the innovations that were encouraged may not contribute to long-term well-being as much as alternative allocations of research investment. Distortions in the pattern of innovation arise from marked discrepancies between social and private returns to innovation.

DISCREPANCIES BETWEEN SOCIAL AND PRIVATE RETURNS AND DISTORTED INCENTIVES We noted in earlier chapters that the social return to an innovation is simply the fact that the product arrives on the market *earlier* than otherwise would have been the case. The private returns are the (incremental) rents accruing to the innovator. Much innovation is directed at seeking and preserving rents rather than enhancing societal well-being. Not only may social and private returns have little relationship with each other, stronger IPR regimes may exacerbate the discrepancies and, therefore, the misallocation of scarce research resources. The following illustrate the nature of the disparities.

Enhancing Market Power Many patents (and the research behind them) are focused not so much on producing a product that is better, valued more by consumers, or cheaper, but rather on enhancing market power, e.g., by extending market dominance. One form this takes is called "evergreening," where a patent holder makes what are fairly obvious slight improvements in the product (drug) to extend the patent and, thus, the firm's market dominance. For instance, a pharmaceutical company, toward the end of the lifetime of a patent, introduces and patents a timed-dosage variant of the pill. Because of the patent, no other producer could have done so. And because the timed-release version is preferred, the effective life of the patent is greatly extended.

The pharmaceutical companies have become experts in extending the effective life of their patents, warding off generics. In doing so, they have made extensive use of trade agreements, which impose a variety of further restrictions. One of the most effective and objectionable is called data exclusivity. A generic producer typically must get approval of its drug, showing the drug's safety and efficacy. This should be easy: All it should have to do is to show that it is equivalent to the patented drug, which has been shown to be safe and effective. But the drug companies say that that is, in effect, using their data (even when the data was partially produced or financed by government), and thus is an intrusion on what should be their intellectual property. But not allowing the generic producer to use the data is an effective way of foreclosing their entry, because it would be unethical to test the generic against a placebo when there is a known product which is effective and safe. Many trade agreements between the United States and developing countries include a provision on data exclusivity (see Charlton and Stiglitz 2005).

Patents can, and often are, used as a barrier to entry. Particularly if the patent is broad, or if monopoly power in one area can be leveraged into monopoly power in another (e.g., Microsoft's dominance in the PC operating system allowed it to extend its dominance into applications, such as word processing), the private return may substantially exceed the social return; the social return to such innovations can even be negative.

Circumventing Patents At the same time, patents sometimes give rise to research of limited social value, as others try to innovate around a patent. The social return on a "me-too" invention (designed simply to circumvent an existing patent and to share in the monopoly rents) is zero.[15]

In short, the patent system not only does not reward inventors on the basis of their marginal contributions, but sets up a distorted set of incentives for innovations, where inventive activity is often directed first at creating market power, and then, by others, at overcoming the artificially created market power.

Patents of Traditional Knowledge There are many other distortions arising from the patent system. There may be relatively low—or even negative—social benefits to allowing patents of traditional knowledge, such as medicinal uses of turmeric. Such patents obviously don't generate new knowledge—the knowledge was already there, and even widely known (except to the patent examiner), but there may be high costs to such patents, since they may impede the use of knowledge. (There are, of course, benefits to transmitting knowledge, e.g., from the sources of traditional knowledge to elsewhere, but these benefits are usually not rewarded with IPR.)

Socially Unproductive Patent Races The patent is granted to the first to make a discovery (or to apply for a patent), and this can give rise to a race to be the first. In the case of decoding the human genome, there was a well-funded international effort, engaged in a systematic scientific process. But others attempted to short-circuit this process, identifying some genes with large market value and racing to beat the more methodical approach. In the case of Myriad and the genes related to breast cancer, they succeeded. But the very limited social benefits of having this knowledge slightly earlier has been dwarfed by the social costs of the impediments (charges) that have been placed on the use of this knowledge.

Mispricing and Perverse Incentives In the market economy, prices are supposed to guide resource allocations—including the allocation of resources to innovation. But if prices are misaligned, then private returns will not accord with social returns.

The financial sector provides the most obvious example (to which we have alluded on several occasions), where much of the returns from innovation were associated with circumventing regulations (intended to reduce externalities, enhance financial stability, and prevent fraud, predatory lending, and market manipulation). While these financial innovations were highly lucrative (even though they were not protected by patents), they had adverse effects on the economy.

Similarly, there is underinvestment in innovations that protect the environment (e.g., reducing greenhouse gas emissions), simply because there is no market price associated with such emissions. But with high unemployment, especially of unskilled labor, there is overinvestment in innovations to further reduce the demand for unskilled labor—which simply create more unemployment and thereby impose high social costs on society (see chapter 6).[16]

Knowledge Pool One of the most important determinants of the pace of innovation (the increase in living standards) is the set of opportunities—the pool of knowledge that can be drawn upon. Every innovation adds to the pool of knowledge, but with IPR—including the "enclosure of the commons," the patent thicket, and the potential for holdups—every innovation also takes out from the effective *publicly available* pool of knowledge. The net effect is ambiguous. Indeed, earlier in this book, we showed that in plausible circumstances, it could be negative. In that model, the adverse effect on the pool of publicly available knowledge of tougher IPR regimes was so strong that the pace of innovation was reduced.

In this section of the chapter, we have identified the static costs and potential dynamic benefits. Advocates of a strong IPR regime typically underestimate static losses, overestimate dynamic benefits, and disregard "balance." Even as they note that IPR may create monopoly power, they emphasize its temporary nature, ignoring the research (noted earlier) showing that such power can persist and that attempts to maintain monopoly distort resource allocations. Poorly designed IPR regimes may even impede innovation.

Later in this chapter, we will explain that there are alternative ways of financing and incentivizing innovation. First, however, we discuss how reforms in the IPR system may reduce the adverse effects of the IPR system while increasing innovation.

2. Reforming the IPR System[17]

There is not just one IPR system. Details matter, and different countries have chosen different rules. This book is concerned with creating a learning society; some rules are more conducive to creating a learning society than others. The effects can be indirect or direct: Rules that allow for monopolization (e.g., longer and broader monopolies) may reduce innovation because of the adverse effects on competition; the Bayh-Dole bill may have reduced the flow of knowledge within universities and the flow of knowledge into the knowledge commons.

Many details of an IPR regime are critical: what can be patented, the standards used for granting patents, the length and breadth of the patent, restrictions on the patent, how the patent is enforced, and rules governing the granting of the patent. In each dimension, there are complex trade-offs, e.g., between providing incentives for innovation on the one hand and introducing inefficiencies in the dissemination of knowledge and impeding follow-on innovations on the other. One of the reasons that there is a broad consensus against patenting mathematical theorems or other insights from basic research is that the disadvantages of patenting far exceed the advantages. Other patents, such as those for business processes, impose other costs, e.g., in terms of litigation risks.

Some of the recent changes in IPR regimes have not always carefully balanced costs and benefits. The extension of the life of copyrights (seventy years after the death of the author) probably imposed more costs than any benefits from improved incentives.

On the other hand, some recent court decisions have shown an awareness of some of the inefficiencies associated with the prior patent regime. We have already mentioned the eBay decision.

With weak standards of "novelty" and "nonobviousness," patent owners can "evergreen" their patents, thereby extending the patent life. There is a legitimate debate about the optimal length of the life of a patent[18]—but such indirect ways of extending it almost surely have greater costs than any associated innovative benefits. The costs can be particularly high in the case of drugs in developing countries.

There have been some excessively broad patents (for instance, the original U.S. automobile patent); the greater the breadth, obviously, the greater the value of the patent, but the greater the impediments for follow-on inventors.

Other details of the IPR regime can reduce the costs relative to the benefits. As we noted, historically, to get a patent, knowledge had to be disclosed, which meant, in principle, that others could make use of that knowledge for their research. Patent rights can be viewed as an exchange, where the "public" grants a temporary monopoly right, circumscribed, in return for the revelation of information. More recently, some in the software industry have been arguing for stronger intellectual property rights *without* disclosure. (In the end, some may choose not to seek intellectual property protection, preferring to rely on trade secrecy.)

Key Issues in the Design of an IPR System

WHAT CAN BE PATENTED Only certain things can be patented. The applicant is supposed to demonstrate, for instance, a certain standard of novelty. Even then, theorems cannot be patented. Some countries have restricted the granting of patents to *processes* for producing chemicals, not to the molecules themselves. The Supreme Court of the United States has recently rejected the right to patent naturally occurring genes.[19] This is a position taken by many other countries. America's business process patents have been widely criticized as extending the reach of patents too far.

BREADTH OF PROPERTY RIGHTS A common misperception is that, once a patent is granted, the grantee has the right to do anything with it during its life time, an uncircumscribed ability to exercise monopoly rights. Property rights do not, in general, and should not, give the owners of property uncircumscribed rights, and this is especially true for intellectual property rights. Intellectual property rights are not an end in themselves; they are a social construction, a means to an end—to promote societal well-being—which is accomplished through careful definition and design.

Examples in which public interest concerns circumscribe ordinary property rights abound; the general principle is that an owner of property cannot do things with her property that might adversely affect others.

U.S. patent law illustrates these issues by excepting patent protection for government use. Under 28 U.S.C. 1498, the government is authorized to use any patent or copyright, a right which can be extended to any contractor, subcontractor, or employee working for the government. While there is extensive debate about the justification for this, the view taken by the United States Court of Federal Claims in the 1990s is telling. It recognized that the granting of a patent was a limited grant—just as it was limited in time, so too it was limited in use. Government use represents a power reserved to the government when it initially grants the patent: "The government cannot 'take' what it already possesses, the government [has] the absolute power to take a compulsory, non-exclusive license to a patented invention at will."[20] While other courts have challenged this interpretation, the Court of Federal Claims decisions make clear that reasonable people, even in advanced industrial countries, balancing interests and looking at the costs and benefits of stronger intellectual property rights, have come to the conclusion that these rights should be heavily circumscribed.

The question, accordingly, is not whether intellectual property rights need to be circumscribed to advance broader social objectives, but how much and in what manner.

CURTAILING ABUSES OF MARKET POWER THROUGH LIMITATIONS ON INTELLECTUAL PROPERTY RIGHTS Intellectual property rights, by definition, create a monopoly power over the use of knowledge, but this is not a license for monopoly abuse. But what is meant by an *abuse* of monopoly power? And what should be the appropriate remedy? There is a broad consensus that Microsoft overstepped the boundaries by leveraging its market power over operating systems into other arenas. Of particular concern is that in doing so, it may have actually stifled innovation. But while both American and European antitrust authorities have concurred on this, they have proposed different remedies, perhaps partly based on differences in judgments about the "balancing" of static and dynamic effects.[21]

One of the responses to abusive, anticompetitive practices has been to restrict the use of patents, effectively insisting on compulsory licensing, sometimes through forming patent pools. In the consent decree in the case of the antitrust action against AT&T in the 1950s, AT&T had to make its patents available to anybody wanting to use them.[22]

Another proposed reform that has gained favor among some academics is limiting the life of intellectual property protection as a way of limiting abuses, increasing market competitiveness, and spurring innovation. If Microsoft's operating system had only a three-year protection, then it would be spurred to make significant improvements in each subsequent release.[23]

Again, what is clear is that there is no unanimity even among the advanced industrial countries on what appropriate balancing entails. (Emerging markets and developing countries should be particularly wary of monopolization; in certain sectors, because markets are less perfect, the threat of monopolization is greater. As we explained earlier, monopolies, once created, tend to persist.)

COMPULSORY LICENSES Besides the restrictions arising from the threat of excessive monopolization, the two most important instances in which patent rights have been circumscribed have been when there is a threat to public health or in response to global warming. The 1992 Rio Agreement created a framework for addressing problems of climate change by providing for compulsory licenses for obtaining access to technology related to mitigation of emissions. The 1994 TRIPS agreement provided for compulsory licenses for lifesaving medicines. (Effective enforcement of this provision has, however, been difficult, since the United States has repeatedly put enormous pressure on governments that have threatened to issue compulsory licenses not to do so. This has remained true even after a global civil society movement succeeded in getting an international agreement clarifying the rights to issue compulsory licenses in the case of lifesaving medicines. [See Charlton and Stiglitz 2005; Stiglitz 2006a.])

THE PROCESS OF GRANTING AND ENFORCING PATENTS Institutional details matter. Nowhere is that more evident than in the way that patents are granted and enforced. Different countries have approached these issues in different ways. In Europe, there is a process of *opposition*: Those who believe the patent should not be granted have an opportunity to express their views to the patent office *before* a patent is granted. There is no such provision in the United States, exacerbating the bias toward excessive patenting that arises from the very structure of the patent system. Patents, as we have noted, privatize knowledge, but challenging a patent converts what would otherwise be a private good into a public good. Thus, challenging is itself a public good and, as in other arenas, the private sector will underinvest in the provision of this public good.

DESIGNING AN IPR REGIME TO PROMOTE LEARNING The previous paragraphs have identified a large number of central issues in the design of an IPR regime. Choices affect the level of competition in the economy and access to and affordability of medicine—and thereby both the health of citizens and (since in most countries, governments pay a large proportion of medical costs) the government's budget. Our focus in this book, however, is on innovation and learning.

The direct effect of IPR is to impede the flow of knowledge and therefore impede the learning process. We have seen that there is an indirect effect—encouraging secrecy—which can be particularly adverse when IPR is extended to traditionally open institutions like universities. Offsetting these costs allegedly are the stronger incentives for innovation and for the acquisition of information. In the case of poorly designed IPR regimes (and the United States and most other countries do not have well-designed IPR regimes), even that effect may be limited—or worse, innovation may be stifled. And even when there are increased incentives for R & D, research may not be directed in ways which increase standards of living, let alone the pace of innovation.

This section of the chapter has outlined some critical details in the design of the IPR regime; a poorly designed IPR regime *on balance*—taking into account the benefits of improved incentives and all the other adverse effects we have noted—leads to a lower pace of innovation; with appropriate design, the balance between the benefits and costs of IPR can be changed.

But as we noted earlier in the chapter, IPR is only one way of incentivizing and financing R & D and learning. IPR needs to be seen as part of a country's innovation system, and there needs to be a balance between IPR and other approaches. This is the subject of the next section.

3. IPR and the National Innovation System[24]

So far, our discussion has focused on the potential disadvantages of the IPR system—that it inevitably gives rise to static inefficiencies and that excessively strong, poorly designed property rights may actually impede innovation and growth. This brings us naturally to the critical question: Are there better alternatives to IPR for producing and financing knowledge?

Alternatives to IPR

Advocates of stronger intellectual property rights give the impression that they are essential for innovation. A moment's reflection should make it obvious that there are many alternative ways to finance and reward innovation. There are, for instance, other ways of appropriating returns from innovation (trade secrets, first-mover advantage) besides patents and copyrights, and in many areas these are highly effective. In fact, as we have already noted, many of the most important advances in knowledge are not protected by intellectual property rights and were not motivated by monetary gains. The basic mathematical advances that provided the underpinnings of the computer and the discovery of DNA which underlies so many advances in modern medicine are but two of a large number of examples.

GOVERNMENT-FUNDED RESEARCH One of the more important alternative ways of financing research is through government support. The United States and other governments finance much of the basic research upon which pharmaceutical companies base their innovations. The Internet, which has spawned myriad innovations since the 1990s, was based on research that was supported, and largely conducted, by U.S. and European governments. The fact that so much of the successful innovation in the United States has occurred in research parks adjacent to universities suggests that these firms are benefiting from knowledge produced in the universities; and university research is, by and large, funded by foundations, government, and university endowments—but not motivated by the search for profit.

More generally, in her book *The Entrepreneurial State*, Mazzucato (2013) has shown persuasively the critical role that the state has played in promoting the large transformative innovations and the smaller innovations, the cumulative effect of both of which has been to create the modern economy.

OPEN SOURCE More recently, the open-source movement has been an important source of innovation. While its original successes were in software, it is now demonstrating its effectiveness in other arenas, such as biotechnology (see Henry and Stiglitz 2010; Hertel, Krishnan, and Slaughter 2003; Lerner and Tirole 2002; Weber 2005). It highlights and strengthens the collaborative nature of research that is the hallmark of academia, and the open architecture facilitates follow-on research—in contrast to the patent system, which closes it down or at least makes it more difficult. As in academia, in some instances nonpecuniary returns play a crucial role in motivating research; in other cases, firms have found a variety of ways of appropriating returns, e.g., through the sale of services or tailoring software based on open source to the needs of particular clients.

PRIZES The prize system represents another alternative to the patent system. This entails giving a prize to whomever comes up with an innovation, or at least those innovations that meet announced objectives.[25] For instance, the person who finds a cure or a vaccine for AIDS or malaria would get a big prize. Someone who comes up with a drug with slightly different side effects than existing drugs (but which is otherwise no more effective) might get a small prize. The size of the prize is calibrated by the magnitude of the contribution.

The idea is an old one.[26] The U.K.'s Royal Society for the Encouragement of Arts, Manufactures and Commerce has been advocating and using prizes to incentivize the development of needed technologies for more than a century. For instance, an alternative was needed for chimney sweeps, those small, underfed boys who used to be sent down chimneys. It was not good for their health, but not cleaning chimneys meant increasing the risk of fire, with serious consequences. So the Royal Society offered a prize to anybody who invented a mechanical way of cleaning chimneys. The prize provided an incentive—and it worked.[27] A patent system might also have motivated the development of a mechanical device (though it did not), but if it had, there would have been a problem: The owner of the patent would have wanted to maximize the return on the innovation by charging a high fee for its use. That would mean that only rich families could have afforded to use the mechanical device, and young boys' lives would have continued to be put at risk. With the prize system, everyone could benefit from this socially important innovation.

The current patent system can, of course, be viewed as a prize system. It is an inefficient one, however, because the "prize" is a grant of monopoly power, and with monopoly power there are incentives to restrict the use of the knowledge. One of the characteristics of a desirable innovation system is that the ideas and innovations, once developed, be widely used and disseminated; the patent system is designed to restrict the use of knowledge. With the prize system, the competitive market ensures efficient dissemination; giving licenses to a large number of people uses the force of competition to drive down the price and to increase the use of the knowledge. With both patents and prizes, market forces are used: one is the incentive of a monopoly to restrict knowledge and raise prices; the other is the force of competitive markets to drive down prices and extend the benefit of knowledge widely.[28]

Moreover, the prize system has the advantage of creating fewer incentives to waste money on advertising and to engage in anticompetitive behaviors designed to enhance monopoly profits. Drug companies spend more on advertising and marketing than they do on research. These marketing expenditures are designed to reduce the elasticity of demand, which allows the owner of the patent to raise prices and increase monopoly profits. From a social point of view, these expenditures are dissipative.

The patent system also distorts the pattern of research. Drug companies have insufficient incentives to develop medicines for the diseases that tend to afflict poor people, simply because there is no money in those drugs. One of the widely discussed ideas for addressing this problem is a guaranteed purchase fund, in which the World Bank or the Gates Foundation would guarantee one or two billion dollars for the purchase of the drug to those who develop a vaccine or cure for AIDS, malaria, or some other disease afflicting the developing world. In effect, there would be a certain market. The guarantee of one or two billion dollars for the purchase of the drug would act as a prize, and a sufficiently large guarantee would provide a clear motivation for research. These guaranteed purchase funds, however, would still maintain the inefficiency of the monopoly patent system, unless there was an accompanying commitment that would make the patent accessible to all at reasonable royalties for purchases beyond the guarantee. The discoverer receives the "prize"—the monopoly profits—by charging monopoly prices. The poor, who get the drugs through the guaranteed purchase fund, do not, of course, pay the monopoly price. But the funds are limited, and when they are used up, without such a commitment, a government that wants to provide to its citizens, say, the malaria medicine that has been bought through the guaranteed purchase fund, will have to pay the full monopoly price.

Money spent purchasing this drug at the monopoly price is money that cannot be spent on the country's other health needs or in supporting basic research or education (each of which could have greater benefits in enhancing societal learning). It may be far better to use the money for the guaranteed purchase fund in a way which spurs competition in providing the drug, to offer a prize, or to buy the patent, and to allow anyone willing to pay a limited licensing fee to produce it.

OTHER MARKET-BASED MECHANISMS Of course, even in the absence of these explicit mechanisms for incentivizing research (patents and prizes), firms undertake research and learning. Firms appropriate returns through natural markets using non-IPR mechanisms such as the advantages that arise from being first or trade secrets. Some industries rely upon such mechanisms even when access to the patent system is possible (partially because they are concerned with the disclosure requirements of the patent law).

Finally, we should note that traditional discussions of intellectual property rights have perhaps overemphasized the importance of monetary incentives. As we emphasized in chapter 3, many, if not most, of the most important advances have been otherwise motivated—by a simple quest for knowledge and peer recognition (see, e.g., Dasgupta and David 1994).

A Portfolio Approach to Innovation

We have described briefly alternative ways in which innovation and learning are financed and motivated. Intellectual property rights should be part of an innovation system that also includes open source, prizes, and government-supported research and grants (which are probably the most important component of the innovation system in supporting basic research). Each of these has its strengths and weaknesses.

Any innovation system has to solve the problems of finance, selection (who gets research money), and incentives. There are, in addition, problems of coordination of research efforts. And different systems have associated with them different implications for the dissemination of knowledge and different transaction costs.

Every country should have a portfolio of instruments. The nature of the portfolio will affect the extent to which the country is successful in creating a learning society; it will affect the innovativeness and efficiency of the system—including the uncertainty and transaction costs facing market participants. In our view, too much weight has been assigned to patents in the current portfolio in the United States. Table 12.1 provides a chart of some of the attributes of the alternatives.

SELECTION The first attribute listed is *selection*. One problem facing any innovation system is how to select those to engage in a research project. The advantage of the patent and the prize system, as well as being "open source," is that they are decentralized and based on self-selection. Those who think that they are the best researchers make the decision to undertake the research. They make the investment, risking their own money, in the belief that they have a good chance of winning the prize (the formal prize or the prize of the patent) or of contributing to the advancement of learning. The prize and patent systems have this advantage over government-funded research, in which there is a group of peers (or bureaucrats) deciding on the best researcher. There is obviously also a concern about "capture" of the research-awarding process, e.g., by political or economic interests whose agendas may be separate from or even counter to the advancement of science and technology.

FINANCE With respect to *finance*, the patent system is the worst of the systems. We can think of IPR as a method of funding research—a highly distortionary method. Price exceeds marginal cost by a considerable amount, for a limited period of time, and the resulting monopoly profit not only provides the incentives but also the resources for innovation. This gap between price and marginal cost can be thought of as a tax, part of the proceeds of which are used to finance research. In the last seventy-five years, there has been considerable research into the optimal way of raising (tax) revenues. The implicit taxation of IPR (even if *all* the proceeds were devoted to R & D) is *not* an optimal way of raising revenues. Its principal virtue is that it is a benefit tax; that

TABLE 12.1
Comparing Alternative Systems

Attribute	Innovation System				
	Patent	Prize	Government-Funded Research	Open Source	Non-IPR Market Appropriation
Selection	Decentralized, self-selection. Lacks coordination.	Decentralized, self-selection. Lacks coordination.	Bureaucratic. More coordination possible.	Decentralized, self-selection. Sometimes "self" coordination.	Decentralized, self-selection. Lacks coordination.
Finance (tax)	Highly distortionary and inequitable.	Can be less distortionary and more equitable.	Most efficient.	May be underfinanced. Foundations, government, by-product of other activities.	Likely to be less distortionary than patent.
Risk	Litigation risk.	Less risk.	Least risk.	Limited.	Limited.
Innovation incentives	Strong but distorted.	Strong, less distorted. Requires well-defined objectives.	Strong nonmonetary incentives.	Strong, often nonpecuniary.	Strong, less distorted.
Dissemination incentive	Limited—monopoly.	Strong—competitive markets.	Strong.	Strong.	Limited—returns depend on secrecy.
Transaction costs	High.	Lower.	Lower.	Low.	Low.

is, only those who benefit from the innovation pay for it. But in most arenas, we do not employ benefit taxes, largely because the additional distortions associated with such taxes are generally not viewed as worth the slight gain in "equity." In the area of lifesaving drugs, such an argument is even more compelling, because typically those who need the drugs are already suffering from having a life-threatening disease. In these circumstances, levying a "benefit tax" by forcing them to pay for the drugs to keep them alive may be viewed as not only inefficient but inequitable. In the case of drugs, the "monopoly tax" is an inefficient way of funding research for yet another reason: A large fraction of the revenue does not reach its target—it is spent on marketing and advertising, rather than on research.

Within the United States, the distortions in consumption associated with monopoly pricing as a basis of research funding for medicines may be limited for those who have health insurance or are covered by Medicare, because so much of the funding for health care, including drugs, is from third parties, and there may be accordingly little price elasticity. Internationally, however, this is not true. In all countries with governments picking up a significant fraction of the cost of medicines, monopoly pricing represents a transfer from ordinary taxpayers to the pharmaceutical companies. And even in the United States, there are large distortions (including in patterns of research) arising out of rent-seeking—in the quest to garner monopolistic returns. But especially in developing countries (and even in the United States and the few other developed countries where governments do not guarantee access to lifesaving medicines), high prices effectively deny access to lifesaving drugs for large numbers of people. Everywhere, for the poor without insurance, monopoly pricing may result in their not getting medicines that they need—they especially bear the burden of monopoly. More generally, the in-effect benefit tax is regressive. A more equitable system of financing would be progressive, and those more able to pay would pay more—and indeed a larger share of their income. Indeed, one can argue that in countries where individuals have to pay a larger share of their drug costs, not charging prices above marginal cost for lifesaving drugs may be a desirable way to provide assistance to poorer people who have the misfortune of suffering from disease. A system of direct payment for the underlying research combined with marginal-cost pricing would make what is going on more transparent, would be a more equitable system of finance, and would lead to better resource allocations.[29]

In short, the patent system is highly distortionary and inequitable in the way in which funds to support research are raised—by charging monopoly prices, e.g., on the sick.

INCENTIVES The main purported advantage of the patent system is that it provides good incentives. Innovation incentives are strong in the patent system, but they are distorted, whereas the prize system can provide equivalent incentives that are less distorted. As we have noted, incentives are distorted under the patent system because there are also incentives to engage in research to innovate around a patent, to spend money in ways that extend the effective life of the patent, to develop a holdup patent, to enclose the knowledge commons, and to extend and to enhance market power. These innovation distortions are in addition to the other market distortions, such as those associated with marketing expenditures and practices attempting to make demand curves less elastic or otherwise to extend and enhance market power.

Moreover, earlier chapters have explained the impact on why innovation of the patent system may actually be very limited—in fact the patent system may be adverse to innovation: The patent system helps create monopolies, and monopoly incentives may be attenuated. Moreover, the patent thicket and holdups have adverse effects on innovation, as they are the attempts by an existing monopolist to foreclose innovative entry of others. And patents create significant impediments for follow-on research.

RISK AND COORDINATION Research is inherently risky: it is an exploration into the unknown. There is, however, a big difference in the nature of the *risk* faced by researchers operating in the difficult systems. Moreover, one of the disadvantages of both the patent and the prize systems is the lack of *coordination*. From a societal point of view, there is a risk of excessive duplication.[30] The lack of coordination increases the cost of doing research.

One of the risks that each researcher faces is not knowing how many other people are engaged in that research. This increases the risk that someone else will make the discovery first, and thus get the patent or prize. Government-funded research can be more coordinated.

Coordination is important with respect to not only the number of researchers or research projects but also the portfolio of research projects. The optimal portfolio takes into account the marginal contribution of each project, given the other projects undertaken—the increased likelihood of a discovery, or that the discovery will occur sooner than it otherwise would have occurred. (See the discussion in chapter 6.)

We have already noted another important source of risk associated with the patent system: that of costly litigation, with uncertain outcomes.

Thus, with respect to risk and coordination, the patent system is the worst, and the government-funded system is the best, because it has the advantages of paying for the input rather than the output. That is to say, researchers get money for their time and other resources spent doing the research, whereas in the prize and the patent systems researchers are rewarded only if their research is successful—and successful before their rivals.

The patent system imposes more risk than the prize system, because under the patent system, there is an additional source of uncertainty: the value of the "prize," which depends on the magnitude of the monopoly rents which the winner of the patent can extract.

One of the reasons that risk is important is that in equilibrium consumers have to pay for the risk borne by researchers. People and firms are risk averse,[31] and if they must bear risk, they have to be compensated for doing so. The patent system makes society bear the cost of that risk in an inefficient way. Under the government-financed research system, not only is risk lower, it is shared by society in a more efficient way.

TRANSACTION COSTS The transaction costs associated with running IPR regimes are very high. Some have claimed that in the United States there is more spending on IPR lawyers (to obtain patents, to pursue patent claims, and to defend against patent claims made by others) than there is on research. While getting precise numbers is difficult if not impossible, it is clear that the transaction costs are considerable, and greater than in the other ways by which innovation can be funded and incentivized.

DISSEMINATION Widespread access to and dissemination of knowledge is one of the most important attributes in assessing the impact of alternative mechanisms within the innovation portfolio (recall that knowledge is a public good, in the sense that there is no marginal cost in an additional individual or firm making use of that knowledge), and here the patent system is particularly deficient. It is designed to impede access to knowledge. Lack of access to knowledge not only leads to a static inefficiency but impedes further innovation. It can be a major impediment to creating a learning society. The hallmark of the alternative mechanisms (open source, government-funded research, the prize system) is that knowledge is made available for free or for a limited licensing fee.

This discussion has made clear that, on most accounts, the prize system dominates the patent system. The prize system provides high returns to innovators who develop innovations of high social value. (Indeed, the prize can be better aligned with social benefits than can the random prize associated with the award of monopoly rights through the patent system.) And then the benefits of this knowledge can be more widely disseminated through the use of competitive production. Moreover, the prize system avoids not just the static but also the dynamic distortions associated with monopoly—including the incentive to impede follow-on innovations (that would reduce the monopoly rents of the patent holder).

But the prize system has one limitation: It does not work when the objective is not well-defined. (There are, however, many areas, such as health, energy conservation, and carbon emissions reductions, in which there are well-defined objectives.) That is why the prize system will never replace the patent system.

At the same time, in basic research—the foundation on which everything else is built—government-funded research will continue to remain at the core of the innovation system. No one has proposed otherwise: The costs of restricting the usage of knowledge associated with the patent system far outweigh any purported benefits. The debate today revolves only around applied research, which entails translating the knowledge acquired in basic research into applications.

4. Intellectual Property and Economic Development[32]

Throughout this book, we have emphasized that what separates developed from less-developed countries is a gap in knowledge—not just a gap in resources. The most successful countries have been the most effective in reducing that gap quickly. Intellectual property rights (and especially a poorly designed IPR system) may impede both access to knowledge and the creation of a learning society in these countries.

We have explained that any IPR regime involves trade-offs between static inefficiency and dynamic gains—any restriction on the use of knowledge induces a static inefficiency, even more so if it gives rise to monopoly power—but that the dynamic gains from the current IPR system may be limited, both because of abuse of monopoly power and the problems posed by the difficulties of defining boundaries precisely, giving rise to patent thickets. Holdups use the patent system to divert to themselves profits that would otherwise accrue to the real innovators. Risk is increased as a result of both litigation and a lack of coordination. And most importantly, impediments in access to knowledge slow down follow-on research. But even if there were significant dynamic benefits, the optimal trade-off for a developing country is different from that for a developed country. Developing countries have much to benefit from learning from those in the developed countries. There can be direct learning benefits, but also, as we have emphasized, indirect benefits, as learning in one industry or one firm spills over to others.

That is why there is a need for a development-oriented intellectual property regime. The intellectual property regime which is appropriate for the advanced industrial countries will not be appropriate for developing countries or emerging markets.

On October 4, 2004, the General Assembly of the World Intellectual Property Organization (WIPO) called for such a regime. Such a regime would begin by asking, How can one more effectively increase sustainable learning by developing countries? (See Stiglitz 2004.)

There were two critical ideas in the WIPO resolution. It recognized that intellectual property "is not an end in itself."[33] And it reiterated WIPO's mission to "promote creative intellectual activity" and "the transfer of technology to developing countries." The new development agenda calls for ascertaining how different intellectual property regimes affect developing countries.

Both the design of the IPR regime and the broader national innovation system have to reflect the differences in circumstances. That is why the attempt at excessive harmonization under TRIPS (the Agreement on Trade Related Aspects of Intellectual Property Rights that was adopted as part of the Uruguay Round trade negotiations) was so misguided.[34]

For an emerging market, access to knowledge is essential for its future growth. Intellectual property should not be used as an impediment to its development. The liability system (in effect, granting a compulsory license at fair compensation for the use of knowledge) is one way of ensuring access to knowledge. Even more modest reforms, such as allowing the use of intellectual property so long as there is a challenge (with appropriate compensation paid if the patent is upheld) would be preferable to the existing system (see, e.g., Lewis and Reichman 2005; Shapiro 2007).

While it may eventually be possible to devise simple rules for judging when, for instance, a compulsory license should be granted, intellectual property rights, especially in developing countries, are at an early stage of development. Simplicity—and the limited capacity of developing countries to engage in expensive litigation—argues that there should be strong presumptions in favor of limiting intellectual property rights when there is an apparent health, competition, or developmental objective. That is, the burden of proof should be placed on the original holder of the patent that there is not a legitimate health, competition, or developmental objective.[35]

Learning in developing countries occurs through several channels, some of which we have discussed in earlier chapters: mobility of people, open-source forms of knowledge dissemination, investment goods, imitation and reverse engineering, formal licensing, intrafirm technology transfer (e.g., through multinational corporations), foreign direct investment, and the acquisition of firms with technology.[36] A few developing countries (notably China) have begun to undertake significant investments in R & D and made extensive use of the patent system. Advocates of stronger IPR have argued that strengthening the patent system would enhance innovation and learning and would increase the willingness of firms from the advanced countries to undertake research, or even move advanced production, which could be the basis of learning, to developing countries. But they ignore the fact that the patent system would impede learning through many of the channels through which learning has typically occurred and that these channels have historically played a critical role. The implication is that stronger (and especially poorly designed) intellectual property rights may have an especially adverse effect on learning and innovation in developing countries.

5. Concluding Comments

Intellectual property regimes are supposed to encourage innovation by providing incentives to do research, enhancing the ability of innovators to appropriate the returns. But intellectual property interferes with the dissemination and transmission of knowledge and encourages secrecy, which impedes learning. Increasingly, there is an awareness of these and other adverse effects, especially for developing countries, of intellectual property regimes, as developed in the advanced industrial countries. Knowledge is the most important input into the production of knowledge, and by restricting the availability of knowledge, IPR inhibits the production of further knowledge (learning). The patent system gives rise to monopoly power; monopolies restrict production, thereby reducing incentives to innovate. The patent system can give rise to a patent thicket, a complex web of patents, exposing any innovator to the risk of litigation and holdup.

There are two implications of this analysis. The first is that, given the critical role of closing the knowledge gap for successful development, the appropriate intellectual property regime for developing countries and emerging markets is likely to be markedly different from that appropriate for the advanced industrial countries. In this area, even more than in others, one-size-fits-all policies are inappropriate.

Second, there are alternative ways of designing an innovation system, with greater emphasis on prizes and on open source. Patents will play a role, but the details of the patent system matter: a good patent system, for instance, has to pay more attention to disclosure, to problems of holdup, and to designing better systems of challenging patents.

The concept of "creating a learning society" provides a prism through which all the policies and institutions of a society need to be examined. This is especially so in the case of policies and institutions that are *supposed* to promote innovation and learning.

This chapter draws heavily upon Stiglitz (2004, 2006a, 2008b, 2013a); Henry and Stiglitz (2010); and Dosi and Stiglitz (forthcoming). There is a large literature on the subject, which we cannot do justice to in this brief discussion. See the references in the above cited papers, as well as the recent books on intellectual property by Perleman (2002); Jaffe and Lerner (2004); Lundvall (2010); Fink and Maskus (2005); Jaffe, Lerner, and Stern (annual); and Cimoli et al. (2014). See also Grandstrand (2005).

CHAPTER THIRTEEN

Social Transformation and the Creation of a Learning Society

WHILE WE have been discussing the *economics* of learning and creating a learning economy, that subject cannot be separated from broader aspects of societal transformation. Much of this book, for instance, has focused on policies that change sectoral composition in ways that would promote learning. But at the root of success is changing mindsets. Change has to be viewed as both possible and desirable, and there has to be an understanding that underneath change is *learning*.

In many ways, understanding how to change mindsets is more difficult than coming to an understanding of what economic policies would facilitate learning. But in creating a learning society, the two are inexorably linked. We want not just to identify what policies might lead to creating a learning society, but also to get them adopted. That requires political systems and mindsets that recognize the virtues of creating a learning society.

The neoclassical model ignores learning, not only paying no attention to the importance of the allocation of resources to learning and research and development, but also assuming that all firms employ best practices—so they have nothing to learn. Not surprisingly, the neoclassical model was not helpful in understanding what is entailed in the creation of a learning economy. Worse still, we have emphasized that the policies that are based on that model are often counterproductive—they impede learning. So too, the neoclassical model, which assumes beliefs and preferences are fixed, is not helpful in understanding changing mindsets.

As we noted in chapter 3, it is not an accident that the change in mindsets that we associate with the Enlightenment was closely associated with the changes in technology that had brought about the dramatic changes in living standards that mark the past two hundred years. And yet, as important as the Enlightenment was to creating the modern economy and society, even in the advanced industrial countries, there are strong forces countering the Enlightenment—questioning basic tenets of science, including and especially those involving evolution.[1]

In developing countries, these issues are all the more important, as Hirschman emphasized in his writings (see, e.g., Hirschman 1958). For instance, race and caste are social constructs that effectively inhibit the human development of large parts of the population in many parts of the world and impede change. The study of how these constructs get formed and how they change is thus a central part of developmental studies. Similarly, Myrdal (1968), in his studies of South Asia, argued, in effect, that certain social constructs affect behavior and were part of what might be called a dysfunctional *economic* and *social* equilibrium that could persist. But he did not address the question of the mechanisms by which such social constructs are created. And what happened in Asia subsequent to his writing showed that societies could evolve. Myrdal also didn't address the question of the mechanisms by which such social constructs evolve, or collapse.

Nor did either Hirschman or Myrdal ask how we might reconcile such constructions of the developmental process with the usual approaches taken by economists, which, for instance, highlight some notion of rationality.

The objective of this chapter is to provide a very preliminary discussion of these aspects of creating a learning society, amplifying on the earlier discussion in chapter 3, and doing so in ways that touch upon the relationship between the approach taken here and that of the standard economic model. This chapter is divided into two sections, beyond the introduction and conclusion. In the first, we present a general approach to societal evolution of beliefs (including beliefs about change) that provides the foundations of an understanding of both societal rigidities and societal change; it provides insights both into the impediments to creating a "learning society" mentality and into how learning mindsets can be created. The second describes the links between democratic ideology and learning societies.

1. Toward a General Theory of the Social Construction of Beliefs and Societal Transformation

Recent work by Hoff and Stiglitz (2010, 2011) has attempted to construct a general approach to societal evolution that clarifies the critical—and unrealistic—assumptions about individual behavior and cognition that underlie what has become the dominant developmental model within the economics literature. This work, at the same time, provides some insights into why Myrdal's (1968) bleak predictions concerning Asia's prospects could have been so far off the mark. The theory proposed by Hoff and Stiglitz, and discussed more briefly here, centers on *what collective beliefs* (sometimes referred to as ideologies) *are an equilibrium*, and how they change. The analysis that follows is divided into two steps. The first focuses on how, at any moment of time, beliefs are formed and what might be meant by *equilibrium beliefs*; the second addresses how beliefs change.

Equilibrium Beliefs

Our theory of equilibrium beliefs is based on three critical hypotheses:

1. *Individuals' perceptions—how they receive and process information—are affected by individuals' prior beliefs*. Well-documented results in psychology show that individuals recognize and process information that is consistent with their prior beliefs in a way that is different from the way they treat other information.[2] This view, based on strong evidence and referred to as *confirmatory bias*,[3] is markedly different from the dominant view of economists, predicated on "rational expectations," where it is assumed that individuals process all information fully and rationally. Such theories are hard to reconcile with both the psychological evidence and the persistence of differences in human views and beliefs.[4]

There are two important corollaries of this. This hypothesis alone can help explain *equilibrium fictions*. That is, equilibrium can be based on beliefs that (to the individual) seem self-confirming even when, in a fundamental sense, they are wrong (see Hoff and Stiglitz 2010). Moreover, it explains why different individuals can have persistently different beliefs—for each, the data that one person sees and how that person processes it confirm that person's prior beliefs. There is no process of consensus building even about the nature of the world, let alone what might be done to improve it.

In this book, however, we are concerned with explaining not just the persistence of disparate beliefs among individuals, but how some societies come to share certain beliefs that are conducive to learning, creating a learning society. Here, we introduce two further hypotheses.

2. *The information most individuals receive comes from other individuals, and how they assess that information (the weight they assign to it) depends on their prior beliefs and their social connections*. The role of society in determining our perceptions goes even further. The cognitive frames which shape perceptions are largely socially determined.[5] They are what is sometimes called social constructs. But in describing them as social constructs, it is important to recognize, as Hoff and Stiglitz emphasize, that they are not necessarily consciously constructed. By "social construct" we simply refer to anything that is collectively made by people. Collective beliefs can emerge endogenously and spontaneously from individual behaviors.[6]

This is a major difference from the neoclassical perspective.[7] Even the categories into which information is placed are often social constructions. Individuals do not choose their "software" in isolation, but within a social context. From among the infinite set of potentially observable data and the infinite ways in which that data could be processed, individuals choose a finite set of data and process them in particular ways; they are limited by the finite set of *socially* constructed categories that are themselves a part of what are called *ideologies* (or *belief systems*).[8]

These hypotheses help explain why different groups can come to believe different things, and why those differences in beliefs can persist, even when the "reality" that each confronts—the true empirical evidence—is the same.

In the end, we are interested in explaining *behavior* (e.g., individuals' performance or the choices they make, including how they allocate time and the sources of information they seek out). That brings us to our third hypothesis.

3. *Perceptions (beliefs) affect actions (choices)*.[9] Perceptions about being powerless, less productive, or unfairly treated affect behavior. Those who think that they will not do well won't in fact do well. If one believes that change is not possible, or that most changes are for the worse, then one won't undertake actions that facilitate and promote change.

The real power of the theory arises from the interactions of the three hypotheses. Individual behavior is based on beliefs that are more complex (or at least different) in their formation than is reflected in standard Bayesian theories about the determination of individual's subjective probabilities of the occurrence of different states of nature.

As in Rational Expectations models, beliefs affect behavior, which affects outcomes, which affect beliefs. But unlike in a rational expectations model, *beliefs also affect what is perceived* and how information is absorbed and filtered. Biases—at every stage of the formation of beliefs—shape perceptions and widen the set of possible equilibria. If individuals come to believe that a certain group of individuals (members of some caste) are less efficient or productive, information that is consistent with that belief will be absorbed more easily than information that is inconsistent with that. Those from other castes will come to believe that their discriminatory attitudes are not discriminatory, but reflect "reality"—for that is how they perceive it. These beliefs are reinforced because they are held by others in the peer group. Even worse, because those who are discriminated against may come to share such perceptions, it may affect their efforts. To some extent at least, the beliefs become self-fulfilling.

The dependence of performance on perceptions, combined with the earlier hypotheses on confirmatory bias and the social construction of beliefs, means that the set of "fiction" equilibria is widened. There can be multiple equilibria.[10]

Just as beliefs affect *individual* actions and performance, widely held beliefs affect *collective* actions. If there is a widespread belief concerning the importance of education and the effectiveness of public education, it is more likely that there will be collective actions in support of public education. And again, there can be multiple societal equilibria.

There can be a high-learning equilibrium, in which members of society believe that change is possible and that education is an important instrument for bringing about change. Such a society will make public and private investments and adopt policies that sustain a learning society, and the outcomes will confirm their prior beliefs—especially when the salience of possible failures is discounted. But other societies can be trapped into a low-learning equilibrium. If individuals believe that change is not possible, they will not make the investments that make change possible.[11]

The Persistence of Socially Dysfunctional Beliefs and Policies

An understanding of the persistence of dysfunctional beliefs systems and the creation of belief systems that might be more conducive to learning is of no less importance for developed countries than for developing countries. The 2008 crisis provides evidence of the relevance of the ideas that we have presented: Those wedded to the notion that markets were always efficient and stable perceived the crisis markedly differently from those who are more skeptical of these perspectives.

To most observers, the crisis (the bubble and the reckless lending that preceded it) demonstrated that markets were not necessarily efficient and stable and that these market failures could be *very* costly. Those who believe nonetheless that markets are fundamentally efficient found alternative interpretations. In "processing" the vast array of "information" about the economy, their prior beliefs led them to discount information suggestive of market inefficiency and instability. While in 2008, it was virtually impossible to ignore that things had gone badly wrong, they sought to blame government—it was the government's attempt to push poor people into housing that was to blame. They held that view in the face of overwhelming evidence to the contrary.[12] When such arguments couldn't work, free market advocates simply said that it was a once-in-a-hundred-year tsunami, that no theory can be expected to explain such unusual events They failed to recognize that the tsunami was actually *created* by the market.[13]

Such considerations are equally important when it comes to theory. Even general theorems (such as that of Greenwald and Stiglitz [1986, 1988]) showing that markets are essentially never efficient if there are information imperfections and imperfect risk markets (as there always are) are "discounted." Theorems, of course, are propositions that follow logically from the assumptions. For a long time, those who believed that markets are efficient tried to find a logical flaw or to question the assumptions of our analyses. But the framework was essentially that used in standard economics—except that there were imperfections in risk markets and in information. No one could deny the small modification that we had made: that information was imperfect and risk markets incomplete. The only recourse was to dismiss these imperfections as *quantitatively insignificant*. Of course, for those who believed in the efficiency of markets, there was no data on the basis of which they could establish such a result. And, anticipating this kind of response, we had gone further, showing that even small imperfections of information had large effects, fundamentally changing the nature of the equilibrium (see Rothschild and Stiglitz 1976; Stiglitz 2002b). Again, devotees of free markets had a way of "framing" the analysis. Either they dismissed the result as a theoretical curiosity or countered that even if markets were not perfect, government attempts to correct the market failure would make things even worse.

Societal belief systems are so important because of the role they play in shaping policy. Chapter 12 analyzed, for instance, intellectual property rights. Certain belief systems that led to and supported the notion that the stronger the IPR regime the better, led to systems of intellectual property rights that, we suggest, impede the creation of a learning society. Neoclassical ideologies which focused on the static efficiency of the economy led to and supported the notion that governments should not undertake trade and industrial policies that might help create a learning society. Beliefs about the virtues of financial innovations led to and supported economic policies that increased macroeconomic instability and undermined resource allocations that would have better supported a learning society.

We've referred to "societal belief systems," but of course this discussion should have made clear that there can exist at the same time groups (within the same or different countries)—subcultures—who have markedly different beliefs—including beliefs about the desirability of change and the effectiveness of alternative policies or actions in creating a learning society.

Changes in Perceptions and Beliefs

Beliefs do change, and with changes in belief, policies and behavior change also. The Enlightenment represented a change in mindsets, one which was conducive to creating a learning society and which provided the basis for scientific inquiry, the fruits of which have been, as we have noted, fundamental to increases in standards of living. But the changes brought about by the Enlightenment were no less profound in the area of social organization. Beliefs about the sources of authority were changed in a fundamental way.

An understanding of how belief systems change—and the extent to which and how those (like governments) who seek to deliberately change belief systems can do so—is, or should be, a core part of the analysis of economic development and history.

Sometimes change occurs very slowly; at other times, seemingly rapidly. Consider, for instance, that for thousands of years, certain types of differential treatment of women were considered not only acceptable but natural, essentially inevitable; and then, in a span of around a hundred years, such behavior came to be viewed as unacceptable in most parts of the world.

Such changes in societal beliefs are sometimes motivated by changes in economic circumstances. But belief systems often have a life of their own, leading to their own evolution. The two processes are typically intimately intertwined. Historical analyses have often focused on how, for instance, changes in the scarcity value of labor (as a result of a change in technology or a plague) might lead to an attempt by those in power to maintain their economic power by imposing restrictions on workers. Belief systems would then evolve to explain and justify these restrictions.

A set of collective beliefs (ideologies) that serves a society—or some group in society—well under one set of economic circumstances may serve it less well under another. And thus, when circumstances change, there will be "forces" to change the underlying beliefs.

It should be clear that particular belief systems may serve the interests of some groups in society over others. Notions of slavery or caste (notions that members of a particular race or caste are in some way inferior) are advantageous to some groups but obviously are disadvantageous to others.

Because belief systems affect the equilibrium, e.g., by shaping perceptions, elites have a strong incentive to influence people's beliefs. (In contrast, in standard economic equilibrium of the rational expectations variety, this is not relevant—cognitive frames play no role.) But the elites cannot simply "choose" the cognitive frames that work best for themselves (nor can nonelites). The task of choosing cognitive frames and imposing them on others is more complicated and is itself constrained by higher-order beliefs, the "ideologies" to which we referred earlier. Even those in power typically do not control all the determinants of the evolution of beliefs. Cultures are always contested.

We have noted, however, that while economic interests and circumstances help explain the evolution of beliefs—and their failure to change in some circumstances—belief systems can have a life of their own. Whether the writers of the Declaration of Independence intended that the notion that all men are created equal extend to women and slaves may not be clear. Once a notion like that is accepted, however, it is inevitable that the meanings get reexamined and reinterpreted.

The broad ideologies—which define both the categories, the prisms, through which we see the world and what particular beliefs are viewed as acceptable—change, but typically slowly, so slowly that at any moment of time they can be viewed as state variables. An institution (like Jim Crow) may be accepted at one time and not at another.[14] It may be part of an equilibrium at one time, and not at another.[15]

Institutions function because they have legitimacy, because they are accepted. The acceptance and performance of institutions depend not only on economic variables but also on the set of general beliefs about the world.

Incorporating "cognitive frames" (ideologies) as state variables provides part of a general theory of societal change that is markedly different from traditional theories, in which only capital and the distribution of power and wealth are state variables. If beliefs have the profound effects that we have suggested they do, and if at times they change, and at others they do not, then a central part of understanding societal evolution is understanding the dynamics of these changes in beliefs—and the circumstances under which rigidities might arise.

In some ways, the latter task is easier than the former. We've explained the concept of equilibrium fictions; if individuals' beliefs are partially or largely dependent on the beliefs of others with whom they interact, there can be a Nash equilibrium in beliefs that can persist, which discounts new information that might contradict those beliefs. Indeed, a set of beliefs that may have been functional at one time, but is no longer so, can persist after the economics or technology that had led to the adoption of the beliefs has changed.

At the same time, if ideologies change, the equilibrium can change, with little or no change in underlying "fundamentals." Changes in views about gender have had profound economic consequences but are themselves only to a limited extent explained by underlying changes in technology or the economy. (Changes in economics may help explain the increased availability of public education, but the extension of public education to women, and the implications of that for societal change, almost surely were driven as much by beliefs about equality as by economics.)

We turn now to one particular set of beliefs that has had enormous importance in shaping societies in recent decades.

2. Democracy and the Creation of a Learning Society

Ideas concerning human rights and democracy have been among the most important in shaping our economy and society. In the United States and Europe, these ideas eventually led to the abolition of slavery, though there were large groups for whom the continuation of this institutional arrangement was advantageous, and those who opposed it reaped little economic gain from the abolition.

Democratic ideals question authority. When America's Declaration of Independence said, *All men are created equal*, it didn't mean that they were of equal physical or mental capacities, but of equal rights, including the right to put forth their ideas into a competitive marketplace of ideas.

But it is exactly that same frame of mind which is so essential for creating a dynamic, learning economy and society. Democracy and an open society are intrinsically interlinked with a learning economy and society. A more open society generates more ideas, a flow of "mutations," which provides not only excitement but the possibility of dynamic evolution, rather than stasis.

Unfortunately, even if in the long run, a more dynamic society benefits most members of society, in the short run, there can be (and normally will be) losers. And not surprisingly, those who might lose seek to prevent such changes through any means they can. The political process is one way that is often taken. Those who seek to maintain inequalities in wealth and power do so not only through policies (economic, legal, etc.) which perpetuate existing bases of power and wealth, e.g., by creating entry barriers but also through policies which attempt to maintain the legitimacy of these inequities of wealth and power. Media policies (control of the airwaves, right-to-know laws, and so forth) thus become important instruments for shaping public perceptions, and thus public policies. The political processes themselves evolve over time, shaped not just by history but by economics, especially in countries, like the United States, where money has such influence in the political process. Firms have learned that they can partially shape individuals' preferences. Those with wealth have more recently learned how to use such tools to shape perceptions in ways that lead to outcomes in the political process that are more favorable to themselves. Sometimes, this entails creating a less open and transparent society—a more open society might lead people to question the persistent inequities, a more transparent society might expose the nefarious ways by which inequities are maintained. When that happens, the long-term success of the economy may be put into question.

Inclusive Growth

So far, we have emphasized the importance of creating a learning economy and society, suggested that success requires not just an economic transformation but a social transformation, and argued that, over the long run, democratic and open societies will be more dynamic. But, as we have noted, democratic processes can be shaped, and there are incentives on the part of some to maintain existing inequities. Democratic processes can then lead to the antithesis of an open and transparent society.

There is thus at least one more requirement for long-term success: inclusive growth. It is now generally accepted that trickle-down economics doesn't work. Higher GDP does not necessarily mean that all, or even most, benefit. The critique of many of the Washington Consensus policies, though, was not just that they were not pro-poor, that is, that the poor did not share in the benefits. Rather, it was that they were anti-poor. Policies that lead to greater volatility (which arguably capital and financial market liberalization do) are anti-poor. It is the poor that bear the brunt of crises—nowhere evidenced more than in the 2008 crisis (see also Furman and Stiglitz 1999). Policies that lead to higher levels of unemployment are anti-poor. Trade liberalization destroys jobs, so that unless such liberalization is accompanied by measures that lead to job and enterprise creation, it can be anti-poor. That is why it is so important that trade liberalization be accompanied by appropriate financial sector and aid-to-trade measures to ensure that job creation occurs in tandem with job destruction. Markets, on their own, do not ensure this, even in seemingly well-functioning advanced industrial countries.

One of the big advances in development in recent years is that we understand not only that some policies lead to anti-poor or non-pro-poor growth but that we have instruments and policies (from broad policies, like micro-credit, to specific instruments, like more efficient cookstoves) to enhance the likelihood that the poor share in the growth that occurs.

The Political Economy of Inclusiveness and Openness

Our argument for why inclusive growth is so important goes beyond the standard one that it is a waste of a country's most valuable resource, its human talent, to fail to ensure that everyone lives up to his or her abilities. Rather, it is based on political economy, on an analysis of how inequality affects political processes in ways which are adverse to long-term learning and growth and inclusive democracy.[16]

Earlier we argued that government needs to play an important role in any economy, correcting pervasive market failures, but especially in the "creative economy," e.g., financing basic research and providing high-quality education. Moreover, innovation is always risky, and in societies with better systems of social protection, individuals are willing to take more risk. Also, societies (like some in Scandinavia) in which there are stronger social protections are more willing to expose themselves and their citizens to growth-enhancing risks, such as those associated with openness.

Consider, for a moment, a society in which there is little inequality. The only role of the state, then, is to provide collective goods and correct market failures. A consensus can be developed on what that entails—since interests are aligned.

But this is not so in societies in which there are large inequalities. Then interests differ. Liberals may want to use the state to redistribute income. While ostensibly conservative high-income individuals may claim that they are only trying to prevent such redistributions, a more careful look at the policies they advocate often reveals that they entail redistributions toward themselves; at the very least, they entail ensuring that the government does not ask them to contribute too much for the support of the public good and that it does not curtail their activities exploiting the poor and extracting for themselves a disproportionate share of public assets. Distributive battles inevitably rage.

Often, the battle takes the form of an attempt to circumscribe government (e.g., an "independent" central bank that is, in reality, accountable mainly to the financial sector, or budget constraints that severely limit the scope of government activity, even when there are very high return investment opportunities in the public sector).

Many in the United States are concerned that the country has embarked on an adverse dynamic, moving it toward an equilibrium in which there will be greater inequality and, as a consequence, toward a less dynamic economy and society. As social protections erode and public investments weaken, including in education, inequality increases. The rich turn to private education, private parks, private health insurance, etc., even though public provision might be far more efficient. Rather than working to improve the efficiency of the public sector, those who seek to limit the scope of government work to tear down the public sector, to undermine its credibility, knowing that if they succeed, then there will be a broader consensus for limiting the role of government and thus limiting the extent to which the government can engage in redistributive activities, *even if in doing so, the government is limited in its ability to engage in collective wealth enhancement*. As this happens, inequalities increase, confidence in public provision erodes, and the state takes on a less important role. It is problematic to gauge whether, in the end, even those at the top benefit; but what is not questionable is that the vast majority in the society lose out.

A casual look around the world suggests that different societies have taken different courses. The Scandinavian countries, by and large, have limited inequalities, have efficient and large public sectors and high standards of living for the vast majority of their citizens, and have succeeded in creating inclusive dynamic economies and societies. There are important differences among the political parties in these countries, but still, there is a broad consensus about most of the elements of the "social contract." America in more recent years has taken a different course. The image of a society with a high degree of social mobility is belied by the statistics, which suggest that such mobility is less than in many "old" European countries. The consequences for the United States are decreasing standards of living for the majority of citizens combined with increasing social pathologies.

There is not a consensus about whether government in a large country such as the United States can achieve the efficiency and effectiveness in the public sector approaching that achieved by the Scandinavian countries. But a major import of the analysis of this book is that even if it can't, it does not mean that markets should be left to themselves. Rather, it means that the instruments have to be adapted to the capabilities of government—and that efforts should be made to improve those capabilities.

3. Concluding Comments

In this chapter we have touched on a broad terrain. Economics, politics, and society are interconnected. Too often, economists have lost touch with these broader dimensions—though we have also argued that much of the conventional wisdom of economists even missed out on the most important *economic* elements in creating a dynamic and creative society.

In discussing the importance of creating an open, democratic, and inclusive society as necessary conditions for creating a dynamic economy and a learning society, we don't want to underestimate the importance of these as ends in themselves. Creativity, voice, and security are all important ingredients to individual well-being and a sense of dignity. The central message of the International Commission on the Measurement of Economic Performance and Social Progress was that GDP was not a good measure of well-being, and policies which narrowly focused on increasing GDP were misguided (Stiglitz, Sen, and Fitoussi 2010).

We have not devoted much attention to specific policies by which we can create a more open and inclusive democracy, or by which we might transform society, enhancing the culture of learning. Nor have we explained how the interplay between changes in economics, politics, and technology, on the one hand, and the self-evolution of belief systems on the other led to the evolution of the Enlightenment; that is a subject which goes beyond the scope of this book. Nor have we had much to say about why the principles of the Enlightenment and the associated learning mindset have still not become universally accepted, or why, even in countries where they are generally accepted, there are segments of society within which these views are not widely accepted. What we have emphasized, however, is that beliefs (the learning mindset) affect not only individual behavior but also collective actions—including policies which affect the extent to which society learns and the pace of economic progress.

We end with two notes. First, we have stressed the importance of equilibrium fictions, beliefs that persist in spite of evidence to the contrary. Two such fictions that have been persistent are (a) the belief in the efficiency of unfettered markets and (b) the importance of "strong" intellectual property rights. (Even the idea that there is such a thing as an unfettered market—that markets *could* be left to themselves—is a fiction, for markets need rules and regulations, and these rules and regulations have to be agreed to collectively.) The irony is that these beliefs—often held to be at the center of creating a dynamic learning economy—may actually have precisely the opposite effect, especially in developing countries.

Second, while policies reflect societal attitudes, which in turn reflect mindsets of the members of society, they also help shape mindsets. Policies which enhance risk taking may lead to mindsets which are more accepting of risk. Intellectual property regimes which encourage secrecy lead to mindsets which value transparency and openness less.

It is, of course, not just public policies that affect learning and mindsets—so do the decisions made by firms. Firms can decide to cultivate an atmosphere of secrecy, which impedes the flow of knowledge within and between firms, or they can encourage more openness. What matters is not just democracy at the level of the nation, but also democracy within the workplace (see, e.g., Stiglitz 2001a, 2001b). Attitudes that question authority can help create a culture of learning at the level of the firm—and that culture can have society-wide benefits. There are, thus, not just "learning" and "technological" externalities, but "mindset" externalities. While the level of innovation is affected by success in creating a learning environment, some kinds of innovation may be conducive to enhancing a learning environment; others can be adverse. Innovations that strengthen the ability to monitor, for instance, may strengthen hierarchical relationships (see Braverman and Stiglitz 1986).

We have stressed in this book that the decisions taken by firms—including the direction of innovation—are themselves shaped by public policy. As we evaluate the consequences of the various policies discussed in the first twelve chapters of this book, we must be mindful of their long-term effects in shaping our society.

This chapter draws heavily from Stiglitz (2010b) and Hoff and Stiglitz (2010, 2011).

CHAPTER FOURTEEN

Concluding Remarks

IT HAS been more than sixty years since our teacher Robert Solow showed persuasively that most of the increases in standards of living were due to technological progress and learning, and since Kenneth Arrow began the analysis of endogenous learning. If we were to evaluate the impact of their work by the number of citations—and even more, on the scholarly work that their papers inspired—the influence of their path-breaking insights has been staggering.

But at another level, the impact on the evolution of economics has been disappointing. True, everyone speaks today of the innovation economy or the knowledge economy, and there have been important advances (some referred to in this book) in the analysis of, say, patents and patent races, and network externalities, to take but two examples.

But the full implications of their work for the neoclassical model, which was central, for instance, to Solow's analysis, have still not been taken on board. And the implications for policy have been even less absorbed into mainstream thinking. This book can be seen as an attempt to fill that void.

Some forty years ago, the revolution brought about by information economics questioned all the standard results and conclusions. Equilibrium might not exist. When it did exist, it could look markedly different from that depicted by the standard model. (Supply might not equal demand. There could be credit rationing and unemployment. Equilibrium might not be characterized by a single price. There might be price dispersions. Price might systemically exceed marginal cost. Market equilibrium was not in general Pareto efficient.) The impact on both theory and policy was profound.

But information can be thought of as a particular kind of knowledge (see Stiglitz 1975a), and we might have expected that the knowledge revolution would have equally profound effects. The objective of this book has been to show the potential that the economics of learning and innovation has for revolutionizing both economic theory and policy. We have questioned, for instance, some of the basic tools used by economists: If most firms are operating below "best practices," does a production possibilities curve based on the assumption that firms are all efficient—or that their knowledge is fixed—make sense? Is it a useful tool *at all*?

We have shown that comparative advantage needs to be reexamined, especially in light of the increasing mobility of skilled labor and capital: A country's long-term comparative advantage is based in part on its comparative learning capabilities.

We have explained why, in a learning economy, there is no presumption that the market economy, on its own, is efficient—in either a static or dynamic sense. Indeed, the presumption is to the contrary. That means that there are policies that can lead to higher sustained growth. But many of the policies that enhance economy-wide learning are the opposite of those derived from the standard neoclassical model. A focus on short-run allocative efficiency may lead to slower growth. Industrial policies—including interventions in trade—will typically be desirable, and they may even be a permanent part of an economy's policy framework, not just in the early catch-up stage.

We have attempted to provide an analysis of factors that increase a society's learning capabilities and enhance its learning. This analysis provided us with a new theory of boundaries of the firm—different from that of Coase (1937)—focusing not on minimizing societal transaction costs, but rather on maximizing learning, recognizing that knowledge can flow more freely within a firm.

We have stressed the importance of viewing learning from a societal perspective. Learning externalities are pervasive, and it is a mistake not to take them into account. Firms, of course, do not, and that means there is no presumption that the market equilibrium—where the market draws the boundary of the firm—is efficient or maximizes societal learning. There is even less of a presumption that the attempts of firms to impose barriers on what other firms can learn from them lead to an optimum.

Most importantly, our focus on learning has provided a new prism through which virtually every aspect of policy—indeed, every aspect of a country's legal framework—needs to be reexamined. And while this is true for all countries, it is especially true for developing countries.

The construction of simple equilibrium models incorporating learning turns out to be a difficult task, because one has to solve simultaneously for market structure and (investments in) learning and innovation. The traditional way that question has been posed, e.g., whether monopoly or competitive structures are more conducive to innovation, is on that account at least partially misleading. (Of course, government can take actions which foster competition, or it can overlook anticompetitive practices.) Beliefs about future growth and industrial structure, moreover, affect current production, learning, and even industrial concentration. We have shown that there may, in fact, be multiple equilibria—a high-growth equilibrium in which it pays to invest a lot in learning today, and as a result there is more growth, and a low-growth equilibrium.

We have examined policies directly aimed at the structure of the economy (like industrial and trade policies) as well as macro-policy. We have also argued that macro-stability is desirable not just because risk-averse individuals dislike volatility, and not just because with higher volatility there may be a larger output gap (a larger gap, on average, between the economy's potential and actual output) and more inequality,[1] but because macroeconomic volatility creates an environment which is adverse to learning. Whatever benefits might arise from the economic restructuring that economic downturns force upon economies are overshadowed by the costs of the losses in learning and innovation. We have shown how policies like financial and capital market liberalization may not only directly lead to an economic structure which is less conducive to learning, but may also create more macroeconomic volatility, with further adverse effects on learning. We have shown too how standard intellectual property rights may impede learning. And we have discussed possible reforms in the IPR regime which may be more conducive to learning. Even more importantly, we have argued that IPR needs to be seen within a broader context—that of a national innovation system, in which open source, prizes, and public investments in research and learning need to be given more emphasis, and IPR less.

We have shown too how government tax and investment policies can be used to foster a learning economy and society. But the list does not stop here. Alternative monetary policies and institutions, investment treaties, education and technology policies, legal frameworks for corporate governance and bankruptcy, indeed, the entire economic regime needs to be reexamined and reevaluated through the prism of learning. We have observed that even systems of social protection can affect learning: Investments in learning are risky, and in economies with good social protections, because there is better risk mitigation, individuals can undertake more risk. There can be more learning. Anxiety impedes learning, and good systems of social protection reduce anxiety. A corollary of the analysis of chapter 13 is that more democratic workplaces may be more conducive to learning, and labor laws that promote such working conditions may therefore be more conducive to learning.

Although in this short book we have been able to address only a few of the many ways that policies need to be reexamined from a learning perspective, our analysis has overthrown many long-standing presumptions:

- We have shown that there is an *infant-economy argument for protection*. Protecting a learning sector with large externalities (which we have argued is typically the industrial sector) leads to faster growth and improved welfare and standards of living, and can support convergence between developing countries and the more advanced countries. Not providing such protection can lead to stagnation. It may be desirable to provide this protection even if the economy never fully grows up.
- There are strong arguments for using the exchange rate as an instrument to encourage the learning sector, especially in the presence of restrictions on the use of industrial policy; it may pay even for a developing country to have a trade surplus—the benefits from learning outweigh the costs of forgoing consumption or investment. And it may even be desirable to engage in this intervention *forever*.

• Broader industrial policies are desirable. Industrial policies are not about picking winners, but correcting pervasive market failures, and especially the market failures associated with learning. Such policies should go beyond just creating a business-friendly environment. In some idealized models, we have derived general formulae for the optimal intervention.

• Financial and capital market liberalization may have an adverse effect on learning, both because of the resulting weaker flow of funds to firms and sectors where learning and learning externalities are more important, and because such policies undermine learning in the financial sector.

• Foreign direct investment can enhance learning, but the extent to which it does will depend on policies, like requirements concerning domestic procurement and employment, which are often criticized within the standard paradigm and often restricted by trade and investment agreements.

• Stronger intellectual property rights—and especially poorly designed IPR systems (and the IPR system in the United States, for instance, is *not* well designed)—can actually impede learning and the creation of a learning society, because they impede access to knowledge, because they encourage a culture of secrecy, which can be antithetical to the openness that facilitates the creation of a learning society, and because the resulting adverse effects on the knowledge pool, which defines the opportunity set facing innovators, discourages investment in innovation.

Three ideas have been key to our analysis. The first we have already referred to: There is no presumption that markets by themselves are efficient. Indeed, while we have highlighted the importance of learning externalities, we have detailed a host of market failures associated with learning. And we have noted that the spillovers from the expansion of industrial firms are not just technological: There are institutional spillovers (e.g., associated with the creation of a financial and educational system), and the revenues raised from the taxation of industry help finance a host of public investment goods that enhance learning and productivity. (Chapter 11 discussed how public investment should be allocated to maximize societal learning benefits.)

We have noted that markets where learning (innovation) is important are likely to be far from perfectly competitive. The only case where, putting aside various sources of diseconomies of scale, competition is viable in the long run is that where there are perfect learning spillovers. But then each firm will try to free ride on the investments in learning and innovation of others, and there will be an underinvestment in learning.

Indeed, in simple models, the economy converges to a monopoly. (With strong Bertrand competition, in the absence of any offsetting diseconomies of scale or scope, a monopoly may emerge immediately.[2]) We have shown that attempts by Schumpeter and others nonetheless to trumpet the virtues of the market economy have not been totally persuasive. Schumpeter was overly optimistic in emphasizing the temporary nature of monopolies—we have shown that they have the ability and incentives to engage in behaviors which allow their monopoly to persist—with adverse effects both on short-run efficiency and long-run innovation. We have shown that those who, following Schumpeter, suggested that potential competition (competition for the market) was an effective substitute for competition in the market were also wrong. Incumbents can deter entry and maintain high profits. Schumpeterian competition does not lead to an efficient allocation of resources to learning and innovation.

We show that one of Schumpeter's propositions is, in general, correct: An increase in competition may lower innovation. But the level of innovation under limited competition (including monopoly) may well be (indeed, in general will be) less than is socially optimal. There are government interventions that can improve welfare and innovation. Though the relationship between the level of competition and the level of innovation is a complex one (and varies greatly depending on particular assumptions), we have suggested that perhaps the real reason that competition may be important goes beyond the standard model of profit maximizing firms, focusing on issues of agency, managerial capitalism, and corporate governance.

We have argued, furthermore, that the failures of the market relate not just to the level of learning and innovation but also to the direction. Too much effort is expended to get, and maintain, market power. Too much effort is also expended on circumventing existing patents. There are large divergences, even with well-designed patent laws, between social and private returns.

One of the manifestations of these distortions is that there may be excessive efforts at saving labor and insufficient efforts at protecting the environment. The result is that the market equilibrium will be characterized by high levels of unemployment—higher than it would be with appropriate government intervention in the innovation process. This distortion is particularly evident in the midst of recession, with high levels of unemployment of unskilled workers, and yet where significant efforts continue to reduce the need for workers.

The second key idea is that we must view markets and government as complements, as working together. It is not a choice of markets *or* government, but of designing an economic system in which they interact constructively. Indeed, as we have repeatedly emphasized, markets don't exist in a vacuum; governments set the rules of the game, and how those rules are written is one of the key determinants of whether a learning economy and a learning society get created. Government can, for instance, help correct the "market failures" that we have argued are endemic in a learning economy. It can provide education opportunities that enhance individuals' capacity and desire to learn. It can provide a system of social protection that provides individuals with the security needed to undertake the risks associated with new ventures. It can support basic research, which underpins the major advances in technology. And it can help prevent the excesses in financial markets that are systematically associated with macroeconomic volatility.

The third key idea is that the design of a learning society is likely to involve complicated trade-offs. We have focused in particular on the trade-off between static efficiency and learning. Many of the policies discussed earlier (including the infant-economy argument for protection) entail a loss in the short run but a gain in the long run. Stronger intellectual property rights, more secrecy, and restraints on the mobility of workers might (all other things being held constant, including the opportunity set facing firms, the pool of knowledge from which they can draw) provide stronger incentives for investments in learning and innovation. But at the same time, they reduce the flow of knowledge; and because knowledge is the most important input into the production of knowledge, they may actually result *overall* in less learning and innovation, and even less investment in learning.

There are other complex trade-offs. Government can design more finely tuned policies which in principle could, if the government had the requisite information and could avoid the political economy problems associated with vested interests, lead to higher economic growth. But in many societies, the distortions associated with such policies arising from vested interests have had adverse effects. In such societies, broad-based policies, such as those associated with managing the exchange rate, are preferable to more finely tuned policies.

But the more general lesson that emerges from combining our theoretical analysis with historical experience is this: Political economy considerations should not affect whether economies engage in industrial and trade policies to help create a learning economy; they should influence only the choice of instruments. Many countries have learned how to manage these political economy problems, and these successes provide experiences from which other countries can learn.

Indeed, looking over the wide sweep of history, almost every successful economy has, at one time or another, engaged in the kind of industrial and trade policies which our analysis suggests.

Social innovations are no less important than the technological innovations upon which economists traditionally focus: The progress of human society has as much to do with such innovations—including innovations concerning how to manage large organizations and organizational learning and how to promote societal learning more broadly, including through industrial and trade policies—as it has to do with improvements in technology.

So too, there are complex trade-offs within firms. A centralized and more hierarchical structure may facilitate coordination (leading to a better designed portfolio of research projects, avoiding costly duplication). It may also lower the risk of undertaking bad projects—projects with a low probability of return. But such organizational designs may, at the same time, stifle innovation and make it more likely that good projects get rejected (see Sah and Stiglitz 1985, 1986).

Achieving balance between centralization and decentralization is a challenge which has to be confronted at every level—within the firm and within society as a whole. As we have noted, with pervasive externalities associated with learning and imperfections of information, there is a presumption that decentralized markets will be Pareto inefficient.[3]

We observed, however, that there again may be multiple equilibria—none of which are efficient; the economy may get trapped in a bureaucratic equilibrium in which innovation is stifled, or in an excessively free market equilibrium, in which imperfectly competitive firms focus innovation on rent seeking.

While we have emphasized the importance of trade-offs, we have also emphasized that there are a variety of policies which can increase both output in the short run and growth and learning. A poorly designed intellectual property regime combined with ineffective enforcement of antitrust laws could lead to lower output today and lower growth.

Most of this book has employed conventional economic models, for instance, with individuals having well-defined preferences; we have explored the consequences of changing only one assumption—we have assumed that learning (as well as market structure) is endogenous. This book has been an exploration of the profound implications of changing that single assumption. Most of this book has modeled learning in a fairly mechanistic way, keeping within the standard paradigm. Yet, as chapter 13 emphasized, both at the level of the individual and society, what is most important is having a learning mindset. Attitudes toward learning are largely socially determined and are affected obviously by societies' experiences. But what experiences a country is exposed to and how those experiences are perceived are themselves affected by beliefs. There can be equilibrium fictions—belief systems that persist because they seem to be confirmed by the world *as perceived by individuals*. While we have only hinted at how beliefs are formed—why dysfunctional belief systems may persist or why, at some time, they may change rapidly—they are at least in part, and perhaps largely, social constructions. Public policies and individual and firm actions are shaped by these belief systems; but at the same time, public policies also shape beliefs. We have noted the irony that certain prevalent beliefs about what is required to create and maintain a learning society may actually impede creating a learning society—that, for instance, beliefs about the importance of strong intellectual property may actually be leading to less openness, creating a culture that is adverse to learning and innovation.

The issues we raise are relevant for all countries, but perhaps especially so for developing countries, as they struggle to close the knowledge gap that separates them from the more developed countries. The Washington Consensus policies, derived from an excessive reliance on the neoclassical model, paid no attention to learning. In focusing exclusively on static efficiency, these policies may have actually resulted in growth and standards of living that were lower than they otherwise would have been. Fortunately, the most successful countries, especially those in East Asia, paid little attention to these policies. Learning was at the center of their development strategies. But unfortunately, not all countries had a choice: Those dependent on the West for foreign assistance, and especially those in sub-Saharan Africa, had to follow these policies. They experienced as a result low (and often negative) growth and deindustrialization.

Countries might like to pretend that they can avoid the issues that we have raised. Standard policies are complicated enough—why complicate them further by worrying about learning? They might hope that they can avoid, for instance, matters of industrial policy—following the neoliberal doctrines that these are matters to be left to the market. But they cannot. The choice they make in each of the arenas that we have discussed will affect the future growth of the economy.

We end where we began: Increases in standards of living have more to do with learning, the focus of this book, than with allocative efficiency, the subject which has been the preoccupation of economists. That this is so holds out enormous prospects for the well-being of those in the developing world: Accumulating resources is a slow process compared to the speed with which gaps in knowledge can be reduced.

But there is more at stake than just an increase in material standards of living: There are profound differences between a stagnant society and a dynamic one, a society in which individuals are struggling to meet the basic necessities for survival, and a society that enjoys the prosperity that modern technology can provide, enabling individuals to live up to their full potential.

The policies that we have described that can help create a learning economy and society inevitably will shape not just the economy, but society more broadly, for the betterment of both, raising still further living standards now and in the future.

Notes

Preface to the Reader's Edition

1. The report was published by the think tank Wetenschappelijke Raad voor het Regeringsbeleid (WRR) in July 2013.

Introduction

1. Or, more accurately, if there was innovation, it was exogenous, not affected by what market participants did.

1. The Learning Revolution

1. How they make those choices will, of course, have profound effects on *measured* growth, since increasing leisure does not show up in conventionally measured GDP (see Stiglitz, Sen, and Fitoussi 2010). This point was emphasized by Keynes (1930). See Stiglitz (2008d) as well as other chapters in Pecchi and Piga (2008).

2. The difference was referred to as the Solow residual. While technical change accounted for *most* of the residual, there were other factors, including the reallocation of labor from low productivity sectors to high productivity sectors (see Denison 1962).

3. Griliches and Jorgenson's work (1966, 1967), which entailed using alternative calculations of the value of capital, suggested a much smaller role for technical progress. Further problems were identified in the quantification of labor input, as economists attempted to assess the role of human capital

in economic growth (Klenow and Rodríguez-Clare 1997; de la Fuente and Doménech 2006).

4. There was a large literature describing how new technologies were "embodied" in capital goods. See, e.g., Solow (1962b) and the discussion and references cited in Stiglitz and Uzawa (1969).

5. For any doubters: engage in a thought experiment in which primitive farmers accumulated more hoes, or even built more irrigation canals. If that primitive accumulation was *all* that had occurred during the past 200 years, standards of living would be incommensurately lower than they are today.

6. As Solow (1956) pointed out, an increase in the savings rate simply leads to an increase in per capita income, not to a (permanently) higher rate of growth. See the further discussion in the next chapter.

7. See Stiglitz (1998c), which describes development as a "transformation" into a society which recognizes that change is possible, and that learns how to effect such changes.

8. This work includes that of Kaldor (1957, 1961); Kaldor and Mirrlees (1962); Uzawa (1965); Nordhaus (1969a, 1969b); Atkinson and Stiglitz (1969); Inada (1963); and Shell (1967) and the papers contained in that volume. This early research addressed not only the question of the rate of technological progress but also the direction (see Kennedy 1964; Samuelson 1965; Fellner 1961; Drandakis and Phelps 1966; Ahmad 1966; and others).

Of course, economic historians have long sought to explain the rate and direction of innovation (see, e.g., David 1975; and Salter 1966). Hicks (1932) made even earlier contributions to this field. More recent work building on these traditions includes Stiglitz (2006b).

9. This work includes the early work of Dasgupta and Stiglitz (1980a, 1980b) trying to endogenize both market structure and the rate of technological progress, subjecting some of Schumpeter's conjectures to more rigorous analysis. Other work includes that of Gilbert and Newbery (1982). Romer's (1986, 1990) work provided inspiration for much of the later work in this area. See Aghion and Howitt (1998) and Romer (1994) for surveys.

10. Arrow's 1962 papers (1962a, 1962b) are the classic references. Key properties of knowledge and its production (knowledge as a public good, nonconvexities associated with the production of knowledge, inherent capital market and risk market imperfections) are discussed at length, with further references. See, in particular, Stiglitz (1987b; based on a 1978 lecture).

11. We do not, unfortunately, use the framework that we develop in this book to answer two key historical questions: What happened to suddenly change the world, to initiate the process of becoming a "learning society"? And why did this process begin where it did and when it did? A few reflections on these questions are contained in later chapters.

12. Gordon "suggests that it is useful to think of the innovative process as a series of discrete inventions followed by incremental improvements which ultimately tap the full potential of the initial invention" (2012, 2).

13. See Solow 1959, 1962b; Solow et al. 1966; Cass and Stiglitz 1969; and the discussions and references cited in Stiglitz and Uzawal 1969.

14. That is, taking into account the costs of creating markets or obtaining information (Greenwald and Stiglitz 1986, 1988).

15. Because industrial policies were often looked upon disparagingly in the years in which neoliberal economic doctrines predominated, some political leaders have looked for other terms to describe such policies, such as "proactive business policies." Alternatively, they have focused on particular categories of such policies (which typically meet with greater approval), such as export promoting policies. We will stick with the more conventional nomenclature.

16. This perspective was reflected in *Knowledge for Development*, the first World Bank Development Report done during Stiglitz's tenure as chief economist of the World Bank (World Bank 1999; see also Stiglitz 1998c, 1999b).

17. Most clearly articulated by Samuelson (1948). See also Ohlin 1933.

18. For a more general discussion of leapfrogging (in the context of patent races), see Fudenberg et al. (1983).

2. On the Importance of Learning

1. Later discussion will explain why we also do not believe that these differences can be explained by the usual kinds of static inefficiencies, e.g., those associated with distorted incentives.

2. To be sure, there may be problems of identification—when changes in capital are required to put into place changes in technology.

3. We recognize that the magnitude and sources of China's increase in productivity have been the subject of some controversy. For a contrary view, see Young (2003), who estimates productivity growth of only 1.4 percent for the nonagriculture sector from 1978 to 1998. There are many pieces of evidence collaborating the rapid increase in standards of living and output, e.g., trade statistics, and we suggest why studies suggesting that total factor productivity growth was low are unconvincing.

4. Studies that suggest that total factor productivity growth has been low typically ignore the lag structure involved in human capita (see, e.g., Fleisher, Li, and Zhao 2010). Moreover, many of these studies simply assume that the factor shares represent competitive returns; in the East Asian countries (and especially China), there is a presumption that that is not the case.

It is hard to reconcile a real return to investment in excess of 10 percent (or even 5 percent) with the patterns of investment, e.g., heavy investment in low-return infrastructure, or investments in U.S. government bonds, yielding real returns that are low or even negative (though the social returns to reserves may be somewhat higher, especially in the early years, when reserves were smaller; chapter 8 presents an alternative explanation for the accumulation of reserves). Studies that suggest low rates of factor productivity growth implicitly are assuming high rates of return to investment (or that is an implicit implication of their econometric analysis).

5. As we emphasize in chapter 3, incentives are relevant not just for investment and labor supply but also for learning.

6. For a recent discussion, see Zhu (2012), who argues that productivity growth has been central to economic growth since 1978. As Zhu notes, China's capital-output ratio has grown little since 1978. While in the initial period of "reform"—1978 to 1988—productivity growth came from agriculture, but between 1978 and 2007, non-state-sector "productivity growth contributed 2.27 percentage points per year to aggregate productivity growth" (119). As we emphasize later in this chapter, improvements in allocative efficiency result in a one-time gain in productivity, not the persistent improvements that were observed, say, in China. (By contrast, Zhu argues, "Overall, gradual and persistent institutional change and policy reforms that have reduced distortions and improved economic incentives are the main reasons for the productivity growth" [104].)

7. Though some studies do show high levels of TFP (total factor productivity growth) for the manufacturing sector in Eastern European countries, the small size of these sectors meant that there was relatively little impact on aggregate productivity. Brandt et al. (2012) find for 1998 to 2007 (340) "firm-level TFP growth of manufacturing firms averaging 2.85 percent for a gross output production function and 7.96 percent for a value added production function."

8. In some countries, like the Czech Republic, multinationals did successfully bring in best practices. The highly educated labor force facilitated the requisite learning. In several of the seemingly successful countries of Eastern and Central Europe, the 2008 crisis made it evident that at least a significant part of this was related to a real estate boom/bubble.

9. Persistent differences across regions in many countries (such as Italy) are evidence of the deficiencies in the standard explanations focusing on artificial barriers to the movement of goods, services, or factors.

10. See, e.g., Foster, Haltiwanger, and Krizan 2001 for a survey. Hsieh and Klenow (2009) find very large gaps in the marginal products of capital and labor across plants within India and China.

11. This analysis does not explain the sources of the differences in learning ability. For example, is it due to differences in management/culture or differences in investments in learning?

12. The unemployment rate in the United Kingdom soared from 4.7 percent in 1979 to 11.2 percent in 1986 (OECD 2011), even as industry production, for instance, increased by nearly 5 percent over the same period (after recovering from an initial drop in the beginning of the decade). (World Bank data accessed through Google Public Data Explorer, GDO production, constant 2000 U.S.$, disaggregated by sector, http://www.google.com/publicdata/directory, accessed February 26, 2013.)

13. One alternatively might have argued that the strike provided greater incentives for efficiency. But if that were the major explanation, why hadn't management adopted incentive structures to encourage these greater efficiencies, which would have saved enormous amounts of labor? The savings would have provided more than adequate compensation for the additional effort. Moreover, this and similar episodes exhibit hysteresis effects: once the organization has learned how to be more productive, productivity remains at relatively high levels even after the exigency which gave rise to the productivity rise is resolved.

14. A single episode of a productivity increase might be attributable to the removal of a static inefficiency; repeated increases should be seen much more as evidence of episodic learning—including possibly learning about how to remove certain static inefficiencies.

There are often periods of negative productivity change. Such periods reinforce the conclusion that much of the action in productivity occurs well *inside* the production possibilities curve.

15. Interestingly, some of the learning involved learning from foreign firms, e.g., about quality circles and just-in-time production (see, e.g., Nakamura, Sakakibara, and Schroeder 1998).

16. Total investment in the United States held steady at between 18.6 percent and 20.9 percent between 1995 and 2001, beginning at 18.6 percent and ending the period at 19.3 percent. From 1981 to 1994 it ranged from 17.1 percent in 1991 to 22.3 percent in 1984, tending to decrease over the period (see the World Economic Outlook database of the International Monetary Fund, available at http://www.imf.org/external/pubs/ft/weo/2012/02/weodata/index.aspx). Gross expenditures on R & D during the same period increased slightly from 2.5 percent to 2.7 percent; from 1981 to 1994 it ranged from 2.3 percent to 2.8 percent (see the indicators of the National Science Board, available as table 4-19 at http://www.nsf.gov/statistics/seind12/c4/c4s8.htm#top, accessed February 26, 2013).

17. Some of the learning was related to computerization; some of the learning was learning how to exploit differences in costs between, say, the United States and China by constructing a global supply chain.

18. With monopolies in the consumer goods industries, the economy still operates along the production possibilities curve, but not at the point along that curve which maximizes societal welfare. With monopolies or imperfections

in competition in inputs, however, the economy will not operate along the production possibilities curve.

19. We do not explore here those distortions or their interactions with learning and productivity growth. Our 2003 book lays out our interpretation of these macroeconomic disturbances. In Greenwald, Salinger, and Stiglitz (1990), we lay out the links between productivity growth and the business cycle (see also Stiglitz 1994c, 2006b; Greenwald, Levinson, and Stiglitz 1993).

3. A Learning Economy

1. In particular, beginning with the work of Stiglitz (1975a, 1975b), Rothschild and Stiglitz (1976), Akerlof (1970), and Spence (1973).

2. Hayek (1945) called attention to the problem of dispersed information—explaining it was precisely because information was so dispersed that central planning could never work. But somewhat inconsistently, he believed that the price system provided an efficient way of aggregating and transmitting information. As Stiglitz (1994c) has argued, if that were the case, then market socialism would have worked. Hayek never formalized his ideas. Later Chicago economists put forward the efficient markets hypothesis, but they too never created formal models to see whether markets efficiently aggregated and transmitted information. Grossman and Stiglitz (1976, 1980) showed that, in fact, they did not. The crisis of 2008 should have settled any lingering doubts about the efficient markets hypothesis.

3. Though, of course, there were other reasons for the failure of our regulatory system, e.g., related to special-interest politics.

4. The concept of learning to learn was developed in Stiglitz (1987a).

5. In 1977 Paul MacCready won the £50,000 Kremer prize offered by the Royal Aeronautical Society for a human-powered airplane with his Gossamer Condor (see http://aerosociety.com/About-Us/specgroups/Human-Powered/Kremer).

6. See Kanbur (1979) and Kihlstrom and Laffont (1979) for the canonical presentation of entrepreneurship in the setting of occupational choice. Emran and Stiglitz (2009) explain why competitive markets may do a bad job learning about who are good entrepreneurs.

7. Supporting evidence includes Asher (1956) and Alchian (1963) on airframe production, Zimmerman (1982) on nuclear power technologies, Lieberman (1984) on production and investment in chemical process industries, and Hollander (1965) on R & D. More recent studies include those focusing on learning by management, in rayon, in semiconductors, and in fuel-cell technology. See, e.g., Walters and Holling (1990); Jarmin (1994); Dick (1991); Gruber (1998); Argote, Beckman, and Epple (1990); Argote and Epple (1990); Barrios and Strobl (2004); and Schwoon (2008). Thompson (2010) provides a recent survey.

8. Arrow's work also gave rise to an extensive theoretical literature. See, in particular, Spence (1981); Fudenberg and Tirole (1982); Jovanovic and Lach (1989); Malerba (1992); Lieberman (1987); Leahy and Neary (1999); Ghemawat and Spence (1985); Young (1991, 1993); and Dasgupta and Stiglitz (1988a). Some of these papers assume market structures that could not plausibly survive in the long run.

9. Warren Buffett, in his 2001 chairman's letter to Berkshire Hathaway investors, available at http://www.berkshirehathaway.com/2001ar/2001letter.html.

10. By the same token, societies (individuals) can develop capacities for learning how to use their leisure well, or they can enhance their capacities to enjoy consumption goods. (This can be thought of as improving individuals' capacities to translate inputs of time and goods into "enjoyment.") Thus, learning also can have large effects on consumption behavior (see Stiglitz 2008d).

11. See, in particular, the appendix to chapter 4 in the unabridged version of this book.

12. An interesting aspect of a failure to adapt is that the school year in many countries is still related to the agricultural calendar—decades after that sector's decline to but a few percent of the labor force.

13. While we emphasize in this book technological knowledge, which enhances the ability to transform inputs into outputs, at every level there are other forms of "learning," e.g., changes in institutions or changes in beliefs, say about the way the economy or society functions. As we explain later, such changes in beliefs may not be based on an accurate analysis of the world and may in fact be counterproductive in terms of creating a learning society (see, e.g., Hoff and Stiglitz 2010, 2011; and the papers cited there).

14. Famously, Google has had a policy of allowing employees to dedicate 20 percent of their workweek to pursuing independent projects.

15. There is a small, but important, literature on an economy's innovation system (Nelson 2004; Nelson and Winter 1993; Freeman 1987; Lundvall 2010). There is also some writing on the "creative economy" (e.g., Florida 2002). Closer to what we have in mind is the work on, for example, agglomeration externalities. Moretti (2011), in his survey, groups the sources of agglomeration externalities in local markets into three broad bins: thick labor markets, thick markets for intermediate inputs, and knowledge spillovers. We might think of the latter as (in our vocabulary) "learning spillovers." Moretti provides references for the current state of knowledge about learning.

16. Earlier, we noted another aspect of learning—learning the comparative advantages (skills) of different individuals. This is a central function of educational systems and is referred to as "education as a screening device." Because of marked differences between social and private return to such screening, market allocation of resources to such screening are not efficient (see Stiglitz 1975b, 2009).

17. There are other differences between young and old that may affect learning behavior. The major asset of the young is their human capital, and there is considerable uncertainty about the value of their human capital. One rational response (but not the only possible response) to this uncertainty is to increase investment in learning.

18. The new learning may, in fact, make the knowledge of those who are older obsolete. While, in the context of the standard competitive paradigm, individuals take the value of assets (including human capital) as given, in the small-scale microeconomics of the workplace, an increase in knowledge (learning) by one party can affect the value of human capital of others.

19. It is worth noting that similar considerations may have played an important role in explaining some of the differences in the transition from communism to the market between, say, Russia on the one hand and Poland on the other. Russia had large, centralized firms, and these were, for the most part, retained as part of the transition. These firms were dominated by older managers. In contrast, Poland had more medium-sized firms and divided more of its large firms up in the process of the transition, providing a greater role for younger managers (see, e.g., Stiglitz 2002a, 2000c; Ellerman and Stiglitz 2000, 2001).

20. Letter from Isaac Newton to Robert Hooke, February 5, 1676.

21. Many in the developing world believe that the intellectual property regime adopted as part of the Uruguay Round trade negotiations impeded development because it impeded access to knowledge. Subsequently, in 2004, the World Intellectual Property Organization (WIPO) called for a *developmentally oriented intellectual property regime*, one in which access to knowledge for developing countries would be pivotal.

22. Moreover, transportation systems are often centered around national hubs. It is easier to move within a country. Institutions and institutional knowledge are also likely to be local.

23. We again note that not all ideas are actually conducive to learning. As we argue extensively in this book, Washington Consensus policies which spread through globalization may have impeded learning.

24. This has provided one of the rationales for why advocates of trade liberalization, such as Grossman and Helpman (1991), suggest that enhanced trade will lead to more learning. As we explain later, there may be other, more than offsetting, effects.

25. But again, the effects can be ambiguous, as individuals are able to create on the Internet communities of like-minded people, reducing the exposure to new ideas (see Sunstein 2001).

26. The ideas in this paragraph are developed more extensively in Stiglitz (1998c). Chapter 13 (based on Hoff and Stiglitz 2010, 2011) elaborates on the role of cognitive frames and how they are shaped.

27. There were, of course, both technological and institutional changes, but they occurred very slowly. There were slow changes in farming technologies that evolved over time. The new world provided new crops, the use of which spread gradually over the entire world. In their time, feudalism and its end, and slavery and its end, represented important institutional changes. The enclosure movement was another institutional change with profound consequences. Many of the changes in technology and institutions were precipitated by exogenous events, such as the Black Plague (see, e.g., Ruttan and Hayami 1984).

Clearly, the Enlightenment helped create the cognitive mindsets that were conducive to innovation and change. Joel Mokyr (2009) suggests that the reason the Industrial Revolution began in England, rather than somewhere else in Europe, has a lot to do with social mindsets, e.g., the belief in the possibility of progress and social norms of honest dealing among businessmen.

28. This is, of course, a simplification. Conniff (2011) writes that this image of Luddites was due to particularly skillful branding, and that in reality the Luddites were not against machines, but "confined their attacks to manufacturers who used machines in what they called 'a fraudulent and deceitful manner' to get around standard labor practices."

29. According to a recent Pew Poll, a third of Americans do not believe there is solid evidence the earth is warming, and nearly 60 percent do not think that warming is mostly because of human activity (Pew Research, 2012, "More Say There Is Solid Evidence of Global Warming," October 15, http://www.people-press.org/2012/10/15/more-say-there-is-solid-evidence-of-global-warming/). And according to a Gallup poll, nearly 50 percent of Americans do not believe in human evolution (http://www.gallup.com/poll/155003/hold-creationist-view-human-origins.aspx, accessed February 26, 2013). If it is difficult to change people's beliefs about matters on which there is such overwhelming scientific evidence, it should be obvious that beliefs about our social and economic system may persist, even in the face of considerable evidence to the contrary. (Beliefs about markets being efficient and stable provide but one instance. While there was a wealth of theory, empirical evidence, and historical experiences suggesting otherwise before the 2008 crisis, it is remarkable how that crisis left the beliefs of so many adherents of "market fundamentalism" essentially unshaken.)

30. The ideas in this and the following paragraphs represent joint work with Karla Hoff and are elaborated in Hoff and Stiglitz (2010, 2011).

31. But that doesn't fully explain why these ideas are adopted by some individuals and groups and rejected by others. Chapter 13 provides a partial explanation.

32. While there are some immediate and important policy implications of these observations, we are not able to pursue them further in this book.

33. These ideas are developed more extensively in Stiglitz (1995b) and Sah and Stiglitz (1987a, 1987b).

34. They are also examples of ideas that are hard to protect with patents, though in some cases, America's business-process patents attempt to do so.

35. We do not comment here on whether their empirical approach really does capture fully the set of related capabilities. Since their work (see also Hidalgo and Hausmann 2009), alternative approaches to characterizing the product space have been explored (see, e.g., Pietronero, Cristelli, and Tacchella 2013). The effects of an improvement in one sector on other sectors depends not just on the similarity of those sectors, but on the institutional arrangements, e.g., providing scope for exploiting linkages. Thus, the fact that natural resource sectors have traditionally not been closely linked to other sectors may be partly a result of the absence of effective industrial policies and the exploitive relationships often evidenced in that sector.

36. See chapters 4 and 9 for further discussions of why geography matters.

37. This section draws heavily upon Hoff and Stiglitz (2010, 2011).

38. There is now a large literature in behavioral economics (with origins in psychology) based on these ideas, including the role of framing and biases in perceptions (see Ariely 2008; Thaler and Sunstein 2008; Kahneman 2011).

39. See Kindleberger and Aliber (2005). The most recent example is, of course, the real estate bubble that led to the Great Recession of 2008. As Stiglitz (2010b) explains, it is hard to reconcile behavior observed there with any notion of rationality (see also Holt 2009).

40. It is only, however, one determinant of learning, as evidenced by beliefs in evolution. While there is a correlation between beliefs in evolution and education and income, the United States stands out as a country with beliefs in evolution that correspond to those of far poorer and less-educated societies. For example, a survey by the British Council and the market research company Ipsos MORI reports that about 33 percent of Americans "agree the scientific evidence for evolution exists." This is a lower percentage than Argentina, China, India, Mexico, or Russia. See "God or Darwin? The World in Evolution Beliefs," *Guardian*, July 1, 2009, http://www.guardian.co.uk/news/datablog/2009/jul/01/evolution.

41. Bénabou (2008b) and Bénabou and Tirole (2006) see individuals as having the ability to choose their preferences (beliefs) so as to maximize their (meta-) utility. Our emphasis, in contrast, is on the social construction of preferences—where "outside" influences play a central role. The individual does not choose their preferences (beliefs) in isolation.

42. Some of what we are saying here can be expressed in terms of standard Bayesian inference: individuals often hold strong prior beliefs, so strong that new information has little impact on posteriors. The literature

on confirmatory bias suggests that priors may be held with far greater conviction than can be justified.

43. This is what Hoff and Stiglitz (2010) refer to as "uber-ideologies." Gramsci argued that "The claim presented as an essential postulate of historical materialism, that every fluctuation of politics and ideology can be presented and expounded as an immediate expression of the [economic] structure, must be contested in theory as primitive infantilism . . ." (1971: 407).

44. These beliefs are held in spite of overwhelming evidence to the contrary. For instance, the bipartisan National Commission on the Causes of the Financial and Economic Crisis in the United States, 2011, agreed, with one dissent, that government efforts to encourage housing among the poor were not responsible for the crisis. Stiglitz (2010b) presents further evidence: not even the default rates on CRA lending (lending directed at poor communities) was higher than that on other lending. Another example is provided by the electricity shortages that developed in the early years of this century in California. Believers in free markets were quick to blame government regulations, particularly those associated with the environment. The real culprit, it turned out, was Enron's manipulation of the electricity market. When the market was re-regualated to prevent such manipulation, the shortage miraculously disappeared.

45. Many years ago, Tibor Scitovsky (see, e.g., Scitovsky 1986) described the drivers of human behavior—including the quest for excitement. See also Bénabou and Tirole (2003).

46. See, in particular, the important work by Hoff (1997). Later work has elaborated on her ideas (see, e.g., Hausmann and Rodrik 2003). Experiments are an important part of the learning strategies of some firms. Hal Varian (2011), chief economist at Google, reports in a letter to the *Economist*, "Last year at Google the search team ran about 6,000 experiments and implemented around 500 improvements based on those experiments. The ad side of the business had about the same number of experiments and changes. Any time you use Google, you are in many treatment and control groups. The learning from those experiments is fed back into production and the system continuously improves."

47. For an elaboration of this argument, see Emran and Stiglitz (2009).

48. Interestingly, the price system typically doesn't even work in the context of firms interacting with each other, simply because it is too difficult to value each individual patent. Often firms create patent pools, agreements to allow each other to make use of certain patents. (Such patent pools can often serve as effective barriers to entry, making it more difficult for firms that are not part of the agreement to enter the market.)

49. The parties to the contract obviously view these provisions in this way. But contract provisions which may be in the private interests of the contracting parties may not be socially desirable (see Greenwald and Stiglitz 1986).

50. We note again, though, that our approach has questioned the relevance of the standard formulation of the production possibilities curve.

51. Some individuals may have an ability to learn quite generally, while others have developed more focused capacities. A well-structured learning society would recognize these differences.

4. Creating a Learning Firm and a Learning Environment

1. That is, it is hard to write good incentive-compatible innovation contracts. For instance, when a firm fails to produce a promised innovation, it is difficult to establish whether it was because of lack of effort or because of the intrinsic difficulty of the task. Cost-plus contracts, or other contracts designed to share the risk of the unknown costs required to make an innovation, have their own problems (see, e.g., Nalebuff and Stiglitz 1983a).

2. An alleged major disadvantage of firms is that transactions within firms are typically not mediated by prices, with all of the benefits that accrue from the use of a price system. But if the benefits of using prices exceeded the costs, firms presumably could use prices to guide internal resource allocations, and some enterprises do so, at least to some extent.

3. For instance, Sah and Stiglitz (1985, 1986) show that the hierarchical decision making which often characterizes large corporations leads to a greater likelihood of rejecting good projects but a smaller likelihood of accepting bad projects. But they go on to show how committees, "polyarchies," and more complex decision-making structures (e.g., polyarchies of hierarchies) can lead to improved decision making with fewer bad projects accepted and more good projects (see Sah and Stiglitz 1988a, 1988b). Large organizations may similarly encounter problems in choosing successor management teams. Sah and Stiglitz (1991) analyze the problem and show how it may be addressed.

4. For a discussion of convergence in productivity across industrial firms, see Rodrik (2013). His results stand in contrast to those presented in chapter 2.

5. For a brief discussion of Xerox, see Wessel (2012).

6. There are exceptions, including the increase in productivity in the U.S. recession that began in 2008. While there are several explanations of this distinctive aspect of the downturn, one is the increasingly short-sighted behavior of firms which ignores the long-run costs of firing or laying off trained workers. In that case, it will still be true that there will be long-run adverse effects of the downturn on productivity. In the Great Depression, productivity growth

also appears to have been quite high, in part due to important investments made by government (including in transportation; Field 2011).

7. This is, of course, consistent with standard results on unit roots (see Dickey and Fuller 1979; Phillips and Perron 1988).

8. For a more extensive discussion, see Stiglitz (2002a).

9. This can be put slightly differently: With capital (debt and equity rationing), the shadow price of capital often increases dramatically in recessions (see Greenwald, Stiglitz, and Weiss 1984; Greenwald and Stiglitz 2003).

10. To identify the effects of the reduced cash flow on firm investment, including investment in R & D, they focus on two situations where changes in cash flow or net worth might be uncorrelated, or negatively correlated, with future expectations. The first study focused on the automobile industry in the United States, particularly in the aftermath of the oil price shocks. Each of these shocks had strongly adverse effects on sales of American automobiles, particularly since they were not as fuel efficient as foreign cars. On the other hand, assuming that American firms could acquire the technological know-how to construct fuel-efficient cars, these oil price shocks should have increased the level of expenditures on R & D. For the unexpected changes in factor prices meant that, while the industry had gone far along the learning curve for large cars, they were still at the beginning of the learning curve for fuel-efficient cars. But the decrease in cash flow had an immediate and direct negative effect on those R & D expenditures, and those firms that were hit the hardest reduced their expenditures the most. Our econometric study corrected for the effect of future sales expectations, and even taking this into account, the effect of cash flow changes on R & D expenditures was significant.

A second study focused on the airline industry in the aftermath of deregulation, which increased competition, lowered prices, and adversely affected cash flows. But the increased output meant that the return to reducing the cost per passenger mile was increased. The evidence was consistent with a dominant role played by cash constraints: The rate of productivity increases declined after deregulation, and those airlines whose cash positions were more adversely affected had the most marked effect on their rate of productivity increase.

Other studies have corroborated these findings. Hall (1992), based on earlier work (Hall 1990, 1991), shows in a large panel of U.S. manufacturing firms during the 1980s that firms that took on more debt subsequently reduced both investment and R & D. In the approximately 250 firms that increased their debt by at least one-half the book value of the capital stock during one year, the decreased R & D expenditures were large enough to account for a reduction in private industrial R & D spending in the United States of 2.5 percent, about one billion 1982 dollars.

11. These effects may be in evidence even in somewhat milder downturns (see Greenwald and Stiglitz 2003; Filippettia and Archibugia 2010; OECD 2009).

5. Market Structure, Welfare, and Learning

1. In *Verizon v. Trinko*, the Supreme Court referred to collusion as the "supreme evil" of antitrust. *Verizon Communications Inc. v. Law Offices of Curtis V. Trinko, LLP* (02-682) 540 U.S. 398 (2004) 305 F.3d 89.

2. This discussion draws heavily upon Stiglitz (2010c).

3. The imperfections of capital markets go deeper: To get a loan, the innovator has to describe to the creditor his project; but the innovator worries that should he do that, the creditor may be able to steal his idea, or build on his idea to create a still better product. Thus, the struggle to appropriate returns from an idea runs into conflict with the necessity to get funding from others (whom one may not be able to trust).

4. It should be remembered too that the antitrust movement of the progressive era focused as much on the political consequences of trusts—the concentration of power—as on the economic consequences, which got fully explicated only with the development of the modern theory of monopoly.

5. As we noted in the previous chapter, Schumpeter even thought (incorrectly in our view) that recessions could have salutary consequences.

6. That is, of course, now changed. The large increases in inequality over the past quarter century mean that even significant increases in average incomes may not be accompanied by reductions in poverty (see Stiglitz 2012b).

7. And eventually, Google's Chrome. In some ways, this experience is partially consistent with Schumpeter's view: While Microsoft has remained the dominant PC operating system for more than three decades now, its dominance in the browser market was much more short-lived.

8. See also Gilbert and Newbery (1982).

9. See, e.g., Baumol (1982); Baumol, Panzar, and Willig (1982); and Martin (2000). The implication was that the "contestable equilibrium" was the same as a constrained Pareto optimum, where lump sum taxes and subsidies (or, more broadly, cross-sector subsidies) were not feasible and where each enterprise had to at least break even. But even if lump sum taxes and subsidies are not available, government interventions, even if restricted to commodity taxes and subsidies, are still desirable (see Sappington and Stiglitz [1987]).

10. Similar results can arise even in the absence of irrationality. If each potential researcher draws randomly from a sample indicating the likelihood of success (the cost of achieving success), then those who get the most favorable draw will undertake the project. (This will be true even though they realize that it is likely that they have enjoyed a more favorable draw than others. The recognition that this is the case will lead them to have a higher threshold before

undertaking the project. The reasoning here is parallel to that of the winners' curse in auctions.)

11. There is extensive literature modelling the various ex ante and ex post competition affects. See, e.g., Aghion, Akcigit, and Howitt (2013); Greenwald and Stiglitz (2014a–c); Gilbert (2006); Vives (2008); and Stiglitz (2014a–e). To our knowledge, there is no model that incorporates all of these effects.

12. The distortions associated with monopolies in the context of imperfect information are more extensive, as they attempt to engage in price discrimination, extracting as much of the consumer surplus as they can (see Stiglitz 1977; Stiglitz 2009, intro. to part 2).

13. Nalebuff and Stiglitz (1983a, 1983b) have shown that well-designed compensation schemes may base pay on relative performance. Indeed, a well-designed contest with just two contestants may, in some circumstances, be optimal.

Patent races can be viewed as a contest: The first to make the discovery gets the prize. But as we have already noted, patent races don't always work as they should to spur innovation; indeed, as we shall shortly discuss, they may even slow down the innovative process. There are also better ways to spur innovation—even using contests with well-designed prizes. The right to be a monopolist is a prize but one that distorts the economy. See chapter 12 and the references cited therein.

14. A caveat: as we explain, what matters are marginal returns, and marginal returns may be increased even though average returns are reduced.

15. Aghion et al. (2013) have emphasized this effect, which they refer to as the business stealing effect. Actually, rent stealing has long been recognized as a central feature in markets with imperfect competition (see note 21 in this chapter), and it manifests itself in a number of ways. The common pool effect described in the next sub-section can also be viewed as a specific manifestation of rent-stealing.

16. There are many dimensions to a patent system, so that it is not always possible to identify when one patent system is stronger than another. The features of the patent system most relevant for our discussion here are those that affect the ability of a firm to "enclose" knowledge that would otherwise have been in the public domain (see Boyle [2003], Heller [1998], and Heller and Eisenberg [1998]) and those features that force the firm to contribute more to the knowledge pool through disclosure (see chapter 12).

17. Some of the adverse effects associated with, say, the U.S. patent system could be ameliorated by patent reform. Others could not. There is an extensive literature detailing the adverse effects of the patent system on follow-on innovation. A recent dramatic example is provided by the patent on the BRAC genes (which play a critical role in determining the likelihood that a woman

gets breast cancer). Before the U.S. Supreme Court ruled against the patenting of genes, Myriad, the patent holder, suppressed the development of better tests for identifying the presence of the gene.

18. There is a large body of research on each of these topics and several overall assessments of the contribution of the patent system to innovation. See, e.g., Boldrin and Levine (2013), Greenwald and Stiglitz (2014a, 2014c), Heller and Eisenberg (1998), Huang and Murray (2008), Moser (2013), and Williams (2013). On the subject of patent thickets and hold-ups, see, e.g., Shapiro (2001, 2010). It appears that the patent thicket is a particular problem in certain sectors, e.g., software and nano-technology. See European Commission (2008), Clarkson and DeKorte (2006).

Most recently, Goldstone (2014) notes that the attempt by the Wright brothers to inhibit follow-on innovation had a disastrous effect on the development of the American airplane industry—so much so, in fact, that at the onset of World War I, no American plane was good enough to go into combat. The industry developed only after the government insisted on cross-licensing.

19. Chapter 12 emphasizes other adverse effects of the patent system—effects which impede the creation of a learning society—and some of these effects may be worsened with an increase in ex ante or ex post competition. The patent system can encourage secrecy, interfering with the "open architecture" which is the hallmark of successful research within academia.

20. In the case of a linear demand curve, the output of each firm is two-thirds that of the monopolist. For a more extensive discussion, see Vives (2008) and Stiglitz (2014c).

21. Dixit and Stiglitz (1977) showed that the market equilibrium was efficient while there was monopolistic competition with free entry, all firms facing the same fixed costs of entry, symmetry, and constant elasticity of demand curves (and no learning). When these idealized conditions were dropped, the market equilibrium became inefficient. One of the reasons for this inefficiency was that some of the rents accruing to one firm (constituting its incentives for entry) were stolen from other firms and did not represent a net increase in consumer welfare.

22. There is a marked difference, however, between the level of innovation under Bertrand competition and the socially optimal level when research is financed out of a lump sum tax. For when research is financed out of a lump sum tax, the price charged will be that associated with the (lowest) marginal cost of production, rather than the lowest marginal cost of production controlled by *another* firm. Because the price will be lower, the output will be higher, and the value of a cost reduction will be greater.

23. See Aghion, et al 2005. For a more recent survey of this literature, see Aghion, Akcigit,and Howitt (2013).

24. See, e.g., Dasgupta and Stiglitz (1980b, 1988a).

6. The Welfare Economics of Schumpeterian Competition

1. See Greenwald and Stiglitz (1986), who develop the concept of *constrained efficiency*, i.e., taking into account the existing differences in information and the costs of acquiring and producing information. Beginning in the late 1960s, Stiglitz had explored the nature of the inefficiencies which arise when there is imperfect information and incomplete risk markets. Stiglitz (1975b) showed that there could exist Pareto inferior equilibria, and Newbery and Stiglitz (1982) showed that trade restrictions could make everyone in all countries better off. Stiglitz (1972, 1982a) showed that stock market equilibria were in general not (constrained) Pareto efficient—Diamond's (1967) earlier result suggesting that they were rested on the special assumption that there was only one commodity, no bankruptcy, and highly restricted specifications of risk. Our 1986 paper in a sense provided a general formulation that embraced these and other earlier studies. (Arnott, Greenwald, and Stiglitz 1994 provides an alternative general formulation. See also Stiglitz 2009.) We extended our generic 1986 results to search and other models in 1988.

2. This was one of the central points made in Stiglitz's 1974 lecture before the Association of University Teachers of Economics in Manchester, U.K. (Stiglitz 1975a). In November 1978, he elaborated on the problems arising from the public-good nature of knowledge in a lecture to an InterAmerican Development Bank–CEPAL meeting in Buenos Aires (published later as Stiglitz 1987b). Knowledge is a special kind of public good—a global public good, the benefits of which could accrue to anyone in the world. After developing the concept of international public goods in an address to a UN meeting in Vienna (Stiglitz 1995b), Stiglitz (1999a) applied that concept to knowledge.

3. As we commented earlier, and we elaborate on later in this chapter and elsewhere in the book, this is true so long as there are not other market failures. Pervasive rent seeking in some innovation sectors may lead to excessive expenditures on some forms of research.

4. One should, perhaps, not put too much emphasis on the fact that these individuals did not appropriate the full benefits of their innovations. There is little evidence that they would have worked any harder with fuller appropriability. Discussions among economists focus on economic incentives; these may be far from the most important determinants of learning and innovation, as we noted in chapter 3.

5. This result provides a telling criticism of aggregate endogenous growth models that have assumed competition. Only the limiting case of perfect spillovers is consistent with full competition, but then there will be little incentive for engaging in R & D or investing in learning (Romer 1994; Stiglitz 1990). Romer's use of Dixit and Stiglitz (1977) preferences provides a simple parameterization within which one can incorporate long-run dynamics in a model

with imperfect competition. As Dixit and Stiglitz note, however, that utility function has some very special properties, and one should be careful about using that utility function, especially for making welfare assessments, e.g., on the optimality of the number of firms (diversity) in the market equilibrium. Alternative specifications can give markedly different results. See, e.g., Stiglitz (1986b).

6. Given the lower level of production, however, the level of investment in learning/R & D may be optimal. When we say that there is less learning under monopoly, we mean *less learning than there would be in a first-best situation where the level of output was optimal and the level of learning reflected that higher level of output*. It can also be shown that the level of learning is less than it would be in the second-best situations where there is government intervention, through subsidies, *even when there are costs to raising the taxes required to finance those subsidies*.

7. Optimal learning may involve producing at a loss, necessitating borrowing (see Dasgupta and Stiglitz 1988a).

8. This is an explanation of the high observed average returns to investment in technology (see Council of Economic Advisers 1995).

9. For a more extended discussion of these issues, see Stiglitz (2006a, 2008b, 2013b).

10. It should also be emphasized that sometimes later arrivals to the market can be in an advantageous position because they can learn from the first firm, avoiding, for instance, some of the mistakes that may have harmed its reputation. Sony's Betamax lost out to a VCR technology that arrived later.

11. See the discussion in chapter 12. Boyle (2003, 2008) refers to the patenting of knowledge that was previously in the public domain as the enclosure of the commons.

12. We noted in chapter 4 that this plays an important role in the determination of the boundaries of firms; within a firm, knowledge moves *more* freely than it does across boundaries.

13. There is an analogy to fishing from a common pool: free entry can result in excessive entry, so that in equilibrium the steady-state flow of fish is lower.

14. Strikes and wars represent similar inefficient breakdowns in bargaining (Farrell 1987). See also Shapiro (2010) for a discussion of holdups in the context of patents.

15. These problems are exacerbated by the fact that the "boundaries" of knowledge are often hard to define precisely. This and the more general problem of the patent thicket are discussed more extensively in chapter 12.

16. These problems can be exacerbated by other deficiencies in the market. Compensation schemes that reward individuals on the basis of relative performance encourage "herding" behavior, where individuals do what others are doing (see Nalebuff and Stiglitz 1983a).

17. In chapter 12, we describe in greater detail the problems that arise in the context of what has been called the patent thicket.

18. The discussion of this section borrows from Dasgupta and Stiglitz (1988a).

19. Inappropriately designed intellectual property regimes can actually inhibit innovation. (See the discussion in chapter 12.)

20. Volcker's 2009 comments were widely reported in the news media. See, for example, Pedro da Costa and Kristina Cooke, "Crisis May Be Worse than Depression, Volcker Says," *Reuters*, February 20, 2009. http://uk.reuters.com/article/2009/02/20/usa-economy-volcker-idUKN2029103720090220.

21. For a more extensive discussion, see Stiglitz (2010b).

22. Greenwald and Kahn (2009) have shown that most of the decrease in manufacturing employment, at least prior to 2000, was a result of improvements in technology (rather than globalization).

There is a large literature supporting the view that innovation in the United States has been "skill-biased." See, e.g., Greiner, Rubart, and Semmler (2003); Goldin and Katz (2008); Autor and Dorn (2013); Autor, Levy, and Murnane (2003); Autor, Katz, and Kearney (2008). Still, movements in wages, wage differentials, and factor shares may not be well explained by theories of skill-biased technological changes. See Shierholz, Mishel, and Schmitt (2013).

23. This analysis is based on Stiglitz (2014d), which builds on Stiglitz (2006b), which itself builds on a large literature on factor-biased induced innovation, going back to Ahmad (1966), Drandakis and Phelps (1966), Fellner (1961), Kennedy (1964), and Samuelson (1965), with antecedents in the literature in economic history (e.g., Salter [1966]; Habakkuk [1962]).

24. See J. Hicks (1932) for a discussion of a typology of innovation. We use the terminology of "factor augmenting" technological change, where technology improvements result in, for example, one worker doing what two workers formerly could do. The demand for labor at the old wage will be reduced if the elasticity of substitution is sufficiently low.

25. Stiglitz (2014d) shows that this is true so long as the elasticity of substitution between labor and capital is less than unity. The more general result is that the market's factor bias is not optimal. When the elasticity of substitution is greater than unity, the long run equilibrium may not be stable; too high of a ratio of effective capital to labor ratio results in a larger share of capital, which leads to a high level of capital augmenting technological progress, which leads to an even higher share of capital.

26. There are other interpretations, discussed briefly. It seems, for instance, that at times wages did not fully adjust and that at times, as a result, firms had difficulty hiring workers. This will result in the shadow wage exceeding the market wage.

27. We also note that the dynamics can be markedly different from those of the standard Solow model, where convergence to equilibrium is monotonic. As we showed in the appendix to this chapter that was included in the unabridged version of this book, convergence is oscillatory. We note, however, that the smooth convergence to the steady state in the Solow model is a function of its extreme simplifying assumptions. Other slight modifications (vintage capital, savings depending on the distribution of income) can also lead to more complicated dynamics. See, e.g., Akerlof and Stiglitz (1969); Cass and Stiglitz (1969); or Stiglitz (1967).

28. This is consistent with the evidence on the stagnation of median wages in the United States over a span of more than forty years and a decline in wages of unskilled workers (e.g., workers with only a high school education).

29. For a more extensive discussion of the issues raised, see Stiglitz (2010b).

30. For a more extensive discussion of these ideas, see Stiglitz (1975a, 1994c, 2010a).

31. We noted earlier that the firms that went bankrupt in Korea were not on average less productive than those that did not.

32. Even more so in countries, such as the United States, where those with money have disproportionate weight in the political process.

33. This section is based on Stiglitz (2014e) and includes passages taken from that paper.

34. In 2013, the United States ranked #5, Sweden #12, and Norway #1; in the perhaps more relevant (as a measure of wellbeing) Inequality-Adjusted Human Development Index, the United States ranked #27, Sweden #7, and Norway #1.

35. According to OECD data from 2012. And taking into account some of the measurement problems noted by the Commission on the Measurement of Economic Performance and Social Progress, the United States probably had an even smaller GDP per hour worked.

36. The interpretation of such data are open, of course, to multiple interpretations. The numbers can depend on the degree of vertical integration. (If there are some parts of the production process with higher value added per worker, a country specializing in those stages of production might appear to have higher productivity, even though productivities in *comparable tasks* are identical.) The numbers can also depend on the mix of skilled vs. unskilled workers. (Obviously, a firm or country that chose to use unskilled workers would have a lower productivity *per worker,* but just as high of a total factor productivity.)

Data on GDP per worker (which avoid some of these issues) are equally plagued by multiple interpretations, particularly related to the fact that GDP is not a good measure of economic performance, e.g., because of problems associated with health care and "defense" spending. See, e.g., Stiglitz, Sen, and Fitoussi (2010).

37. And for good reason, related to an assessment of the costs and benefits of patents.

38. Though Mazzucato (2013) persuasively demonstrates the limited role of the venture capital industry in innovation. The venture capital industry is a very small part of the financial sector and was adversely affected by the global financial crisis which was brought on by the dominant part of that sector. Kaplan and Lerner (2010) find that historically venture capital investments in companies represent a remarkably constant 0.15 percent of the total value of the stock market.

39. The Programme for International Student Assessment, administered by the OECD, evaluates fifteen-year-old students' aptitude in reading, mathematics, and science literacy. According to PISA, the U.S. education performs at about the average level of OECD countries overall but lags behind the OECD average in mathematics. (Organisation for Economic Co-operation and Development 2011). The low level of equality of opportunity implies that those born to poor and poorly educated parents are less likely to live up to their potential. See Stiglitz (2012b). These adverse outcomes can be thought of as a natural outcome of the American model of capitalism, which has led to high levels of economic inequality, especially given the manner in which these economic inequalities interact with political processes—leading to low levels of public investments.

40. See Gertner (2012).

41. But note that *none of these schools are for-profit institutions*. They are either not-for-profit or state institutions.

42. This is the essential insight of Domar and Musgrave (1944) and Stiglitz (1969). The details of the tax system affect the extent of risk sharing, and thus the extent to which innovation is encouraged.

43. Earlier, we referred to the important role that social attitudes and mores can play; the Enlightenment was a change in mindset, and that change in mindset was far more important than any change in property rights or incentive structures. So too here; attitudes toward failure can affect an individual's willingness to undertake risks. The determinants of these social attitudes would take us beyond the scope of this paper, but there is a growing body of research emphasizing the role that government policies can play. See the World Bank's *World Development Report 2015*.

44. This is not the only reason that there may be underinvestment in education. Some individuals, particularly from underprivileged families, may not fully appreciate the returns to education. The assumption of fully rational expectations assumed in conventional models is clearly wrong. Most individuals rely on public provision of education at the elementary and secondary level, and there may be underprovision of investments, especially in communities in which there are large numbers of poor individuals, in divided societies where rich individuals have access to private schools.

45. For instance, the gains of the high-frequency traders occur at the expense of others, but this rent-seeking not only uses real resources, it makes markets less informative with adverse effects on the efficiency of resource allocation. See Stiglitz (2014g) and Biais and Woolley (2011).

46. There is a long tradition among economic historians arguing for the innovation benefits of high wages and labor scarcity. See, e.g., Salter (1966), Habakkuk (1962), Sutch (2010), and Wright (1986). For a theoretical discussion, see Acemoglu (2010), Greenwald and Stiglitz (2014a), and Stiglitz (2006b, 2014d).

47. See Stiglitz (2012b) and Piketty (2014) and the references cited there.

48. See, e.g., Hoff and Stiglitz (2004a, 2004b).

49. There are multiple links between inequality and the economic-political equilibrium. More divided societies are less likely to make high-return public investments; the rich seek a weaker state, worrying that it might use its powers to redistribute. See Stiglitz (2012b) and the references cited there. See also Bénabou (1996) for a survey of studies on economic growth and inequality and Ostry, Berg, and Tsangarides (2014) for more recent evidence.

50. We say this with some confidence, having watched closely and participated in decision making in the United States, and especially relevant for this paper, decision making related to innovation, such as the design of intellectual property rights and the level and pattern of expenditures on research. Special interests often dominated. The question of what was good for the progress of science or the advancement of health was given short shrift. For a discussion of some aspects of this, see Stiglitz (2006a).

51. The fact that there are a few countries, such as Korea, that have gone a long way in closing the knowledge/technological gap, is not necessarily inconsistent with this hypothesis. Such countries may have underestimated the cost of closing the gap, may have unusually low time discount factors, or may face distinct circumstances in which the costs of closing the gap are unusually small.

52. There are other models generating non-convergence. In the appendix to Stiglitz (2014e), we discuss the important differences between our model and that of Krugman (1981) and Matsuyama (1992).

53. By the same token, some of the policies of the United States that impede innovation are a result of political processes that reflect the influence of special interests, like the financial sector.

54. In the standard competitive model, where each firm faces a horizontal supply curve of labor of each type, these effects are not likely to rise. But in practice, labor mobility is imperfect. Firms are engaged in a bargaining process with their workers. The nature of technology—which they can shape—affects this bargaining process. Note that this analysis does not require that firms coordinate their actions to increase their bargaining power vis-à-vis workers (though under some circumstances they may in fact do so). Rather, so long as

there is imperfect mobility of workers, it pays each firm to take actions which increase its bargaining power vis-à-vis its workers.

55. See Braverman and Stiglitz (1986) for an analysis of these issues in the context of an agricultural economy.

56. In perfectly competitive labor markets, any nonpecuniary cost would lead to a demand for higher wages and thus would be taken into account by the firm, but this is not so in imperfectly competitive markets.

Moreover, in a world with uncertainty and imperfect information, management may come to believe that technologies that save on labor are profitable. There may be an "equilibrium fiction" in which the evidence, as they see it, confirms those beliefs (Hoff and Stiglitz 2010). This is especially so in managerial capitalism (with agency costs which enable managers to exercise considerable discretion for their own benefit), where managers may value their own time and trouble more than would be the case in an efficient market economy.

Interestingly, such beliefs will, in fact, serve the interests of the managerial/capitalist "class" as a whole, leading to outcomes that are consistent with what they might have wanted to do collusively, though they had no mechanism by which to do so.

57. There is, in this sense, a kind of increasing returns to scale. The more innovators think about how to improve labor efficiency, the better at it they get. This suggests that the innovation frontier, rather than being concave, may in fact be (at least in part) convex.

58. This is one of the points raised by Phelps (2013).

59. East Asia did this as a central part of its development strategy (see Stiglitz 1996; World Bank 1993).

7. Learning in a Closed Economy

1. The precise nature of the distortions is complicated. Dixit and Stiglitz, for instance, suggest that there may be too many firms with high price demand elasticity with low levels of consumer surplus. Stiglitz (1986a) derives conditions under which there will be too many firms (too much product diversity). Here, we need to ascertain the effects of these market biases on the overall pace of learning. For instance, under the conditions in which there are too many firms, each will have a more limited incentive to engage in R & D (relative to the social optimum), so the pace of innovation will be slower than is optimal.

2. In a simple two-period model, the precise condition for monopoly output being greater than the competitive output is

$$1/\eta_M < \delta \nu h_M$$

where η_M is the elasticity of demand for the manufactured good, δ is the pure time discount factor, υ is the ratio of labor supplied in the two periods, and h is the elasticity of the learning curve (the percentage reduction in next period's marginal cost of production as a result of a one percent increase in this period's production). The variable υ itself depends on the labor supply elasticity. A higher labor supply elasticity means that when costs (prices) fall, there is a greater increase in labor supplied. Similar results hold in more general models.

3. There are two other situations which deserve brief mention. The first is where there are not full spillovers, but competition is maintained within the learning sector as a result of diseconomies of scale. To take the extreme case, assume there were no within-sector learning spillovers but only cross-sector learning spillovers. Then each competitive firm would take full account of the benefits of its learning to itself—just as the monopolist would. In that case, the only distortion in the competitive equilibrium arises from the cross-sector externality. As long as there are spillovers, production in the learning sector will be sub-optimal, and a first period consumption subsidy would be desirable.

The other case of interest is that where there is imperfect competition in both sectors. If the degree of monopoly is roughly the same, then both will raise prices relative to marginal costs, but relative prices will be unaffected. Hence, monopoly power won't distort (relative) production of the two goods, but will result in internalization of the within-sector externality, but not of the cross-sector externality. Of course, the reduction in real wages not only has distributional effects, but it will also affect labor supply, and thus learning.

4. $\rho = \delta\, V_y^{t+1}/\mu$, where δ is the pure rate of time discount, V_y^{t+1} is the marginal utility of income (at $t + 1$), and μ is the marginal utility of income *in the public sector* at time t. Thus, ρ is more accurately described as the value of a dollar tomorrow relative to the value of a dollar today within the government's budget constraint.

8. The Infant-Economy Argument for Protection: Trade Policy in a Learning Environment

1. Indeed, in some circles, opposition to free trade would be grounds for taking away one's certification as an economist.

2. Actually, the circumstances in which free trade is welfare enhancing are more restricted than is justified by this widespread presumption. For instance, when there are imperfect risk markets, free trade can actually make all individuals worse off (see Newbery and Stiglitz 1982). For a broader discussion of these issues, see Charlton and Stiglitz (2006, 2012).

3. It is worth noting that in popular discussions, it is often argued that openness leads to more learning and that there are learning benefits associated with trade. While this may be true, this positive learning effect needs to be offset against the effects associated with the structure of production. For reasons set forth in earlier chapters, we believe that the latter effects predominate. For a contrasting view, see Grossman and Helpman (1991), who, however, essentially ignore the effects upon which we focus.

4. There is a long history of the infant-industry argument for protection, dating at least back to the mid-nineteenth-century work of List (1841). For a more extensive discussion, see Chang (2002, 2003); Charlton and Stiglitz (2005); and Stiglitz (2006a). The discussion here borrows heavily from Dasgupta and Stiglitz (1988a).

5. Recall that we are using these terms as metaphors. *Agriculture* includes small-scale rural nonagricultural and craft activities. *Industry* may even include industrial agriculture.

6. This is discussed at greater length in chapter 5 and formalized in earlier chapters in Part Two of the unabridged version of the book.

7. Stiglitz and Weiss (1981) explain why there may be credit rationing with imperfect and asymmetric information.

8. In the case of East Asia, governments used rule-based systems, providing more finance to those firms that had demonstrated prowess in exporting, and especially in areas where there were significant potential technological spillovers.

9. A bank's granting of a loan, in this case, is little different from a government's decision about which researcher to support—except that in the latter case, the government can simultaneously evaluate different research applications, while the bank can only guess at what other researchers are receiving funding. Moreover, the government can assess the marginal social return associated with each project, while the private lender has to judge the expected (private) return of this particular project, i.e., the (average) probability of success times the profits that the firm will get if successful. As explained in earlier chapters, in the case of innovation, expected private returns are not closely aligned with (expected) marginal social returns. With government financing, projects get funded so long as expected marginal social returns are positive. With private financing, projects get funded so long as the bank's expected return is positive. In both cases, those with more learning potential get more access to funds. But the cutoff and the mix of projects can be markedly different. In particular, the government can take into account learning externalities. There is one other critical difference: if a loan goes bad when the private sector provides finance, it is shareholders who bear the cost (unless there is a government holdout); if the government provides the funds, it is taxpayers who bear the costs.

10. One might argue that since patent protection is time limited (though firms have found clever ways of extending the effective life of patents), so too should protection be time limited. The analysis presented in this chapter shows that this may not be correct.

11. See chapters 3, 4, and 12 for a discussion for why geography matters for spillovers.

12. The only assumption that differs from that of earlier chapters is that concerning spillovers over space. We assume that there are perfect spillovers within a country, but no spillovers across borders. Obviously, this is a limiting case, and the more general case can be analyzed as in chapter 8 in the unabridged version of this book.

13. Alternatively, it was argued that if it eventually should develop a comparative advantage in manufacturing, there was no point in anticipating the change. The critical assumption was that technology was exogenous.

14. As we noted in chapter 1, there are difficult problems in ascertaining a country's dynamic comparative advantage.

15. In Arrow's original 1962a paper, learning was related to the level of investment, not the level of production.

16. See Korinek and Servén (2010). Other instruments, such as investment subsidies/tax credits and interest rates, are also relevant. For a discussion of the use of exchange rates (and changes in exchange rate) in the context of the East Asia miracle, see Stiglitz (1996).

17. As we noted in chapter 5, the existence of profits may not suffice to attract actual entry. Entrants care about what the market will look like after entry, and they may believe that after entry, competition would be sufficiently keen that they would make a loss. The incumbent can take entry-deterring actions which reinforce such beliefs. The incumbent can, in particular, undertake sufficient learning that it preempts entry of rivals.

18. It should be clear that in a more competitive marketplace, the sum of profits should be lower.

19. With constraints (and costs) associated with levying taxes, especially in developing countries, protectionism (a hidden tax) may seem preferable. But this lack of transparency is an important argument against protectionism.

20. This is similar to the "rent-stealing" or "business-stealing" effect discussed in chapters 5 and 6. From a global point of view, there may be no social benefits, even if the developing country is better off.

21. Taxpayers in the country bear the costs of the infant-industry subsidy to help create an effective competitor. Government must judge whether such subsidies are worthwhile by assessing the value of the future profits it can glean from the eventual profits of the entrant plus the value of the consumer surplus that accrues to their citizens, ignoring the benefits to citizens of other

countries. (If the monopolized good is an input into production, there can be further benefits from competition: higher profits to the firms that use the good as an input, higher consumer welfare from the lower prices on consumer goods that may result, and higher tax revenues to the government. Such benefits are themselves global public goods, since all producers anywhere benefit. But there are likely to be localized benefits as well, and not just from knowledge spillovers; there can be, for instance, beneficial design interactions between the user and producer of the intermediate products.)

22. In some simple models, it may even be desirable for them to run surpluses in perpetuity.

23. Though one should note there is some evidence of convergence within manufacturing. See Rodrik (2013).

24. See Noman and Stiglitz (2012a, 2012b, 2012c), and the references cited there.

25. That is why empirical studies showing correlations between "openness" and growth are misleading. The fact that some countries that have followed excessively protectionist policies have tried to shut themselves off from the rest of the world has no bearing on the design of optimal policies.

26. This is the main theme of Stiglitz (2006a).

27. At the same time, it is important to recognize that the uncoordinated decisions of different countries is not likely to lead to an efficient global equilibrium. Global coordination is desirable, but will be hard to achieve if advanced countries demand that the global rules be such as to serve their interests.

28. An important aspect of the global financial architecture is the global reserve system. We have criticized existing arrangements, arguing that they contribute to global instability and an insufficiency of global aggregate demand.

29. This section reproduces in part Stiglitz (2014e).

30. The empirical literature on convergence is complex. See De Long (1988); Durlauf and Quah (1999); Dowrick and Nguyen (1989); Barro and Sala-i-Martin (1991); Mankiw, Romer and Weil (1992); Im and Rosenblatt (2013). This literature focuses on countries with the same production functions, assuming that knowledge flows freely across boundaries. It is precisely this question upon which we focus here.

31. Formally, we assume $g^L = f^L(\pi^L, \kappa)$.

32. It should be noted that the two curves may take on different shapes. For instance, what matters for the determination of the optimum is the marginal return to allocating more labor towards manufacturing. If the gap in knowledge is too large, even though there is more to learn, the ability to learn may be lower, so the marginal return will be lower. In this case, the *PM* curve could be upward sloping for large values of κ. There can then exist multiple steady-state equilibria.

9. The Role of Industrial and Trade Policy in Creating a Learning Society

1. See Stiglitz (1998c), which describes development as a "transformation" into a society which recognizes that change is possible and that learns how to effect such.

2. We should be somewhat more careful. Williamson (1990) articulated the Washington Consensus in the context of the policies that were pushed in Latin America, and Africa's structural adjustment policies began well before his formulation of these ideas. But the underlying beliefs about economics and economic policy, which Williamson put so clearly, had long dominated development thinking in the international economic institutions, and the ideas were applied not just in Latin America. Williamson actually distanced himself from some of the ideas that subsequently get identified with the Washington Consensus. He was, for instance, always cautious about capital market liberalization, and capital market liberalization was *not* part of the set of policies he identified as the Washington Consensus (see Stiglitz 2008c; Williamson 2008; as well as other papers in Serra and Stiglitz 2008). For further discussion of the Washington Consensus, see Stiglitz (1998a, 1999c, 2002a).

3. In chapter 11 we will explain why even this prescription may have been misguided, at least from our learning perspective; what matters is *real* stability, as much or even more than *price* stability, and the excessive focus on the latter may have undermined the former (see Stiglitz et al. 2006).

4. It is, perhaps, worth noting that what is viewed as corruption in one society may not be viewed in that way by others. Many outsiders look at the American system of large campaign contributions and revolving doors, which seems to "buy" favorable legislation, as a form of corruption, even if there isn't money stuffed into brown paper envelopes for the politicians themselves.

5. Dixit (2012) has argued that firms from developing countries may have a knowledge advantage in dealing with governments of other developing countries.

6. Herbert Simon emphasized that if there are differences in the performance of public and private enterprises, the differences could not be explained just by differences in incentives, since in both, typically, most individuals work for others and have to be incentivized (see, e.g., Simon 1991). "This examination of authority and organizational identification should help explain how organizations can be highly productive even though the relation between their goals and the material rewards received by employees, if it exists at all, is extremely indirect and tenuous. In particular, it helps explain why careful comparative studies have generally found it hard to identify systematic differences in productivity and efficiency between profit-making, nonprofit, and publicly controlled organizations" (Simon 1991, 39).

7. Rashid (2011, 2012) and Detragiache, Tressel, and Gupta (2008) provide data and an empirical analysis strongly supporting this conclusion. Greenwald and Stiglitz (2003) present the general theory.

8. See, e.g., Shapiro and Stiglitz (1984); Stiglitz (2002b); and the references cited there.

9. It is difficult to track inequality because of data limitations. The *Africa Progress Report* states that twenty-four countries in Africa have Gini coefficients in excess of forty-two, the level in China. It also points out that in a number of cases, recent growth has not been matched by falling poverty—which they attribute to inequality: "In many countries, the pattern of economic growth is reinforcing these inequalities" (Africa Progress Panel 2012, 16).

10. See, e.g., Newbery and Stiglitz (1982), who show that free trade can make everyone worse off (that is, it can be Pareto inferior) when there are imperfect risk markets, because it increases risk.

11. There is a long-standing theoretical presumption that this would be so for advanced countries (Stolper and Samuelson 1941), but there is also evidence that this is so even for developing countries (see Stiglitz 2006a). With trade liberalization often associated with an increase in unemployment, it is not surprising that there are adverse distributional consequences: those at the bottom are most likely to be laid off, and higher unemployment puts downward pressure on wages (see Furman and Stiglitz 1999). The adverse effects of trade liberalization were often exacerbated by simultaneous measures liberalizing financial and capital markets, which contributed to economic volatility (see, e.g., Stiglitz 2008a, 2010e, 2012a). For further discussions of the possible adverse effects of liberalization on inequality, see World Bank (2005); and Topalova (2010).

12. See, e.g., Rodriguez and Rodrik (2001). Wacziarg and Welch (2003) found that roughly half of the countries in their survey experienced zero or even negative changes in growth post-liberalization.

13. A few econometric studies (cross-country regressions) have been particularly influential (see, e.g., Dollar 1992; Sachs and Warner 1995). But while these authors were careful to qualify their results, others have not been.

14. See Charlton and Stiglitz (2006, 2013). They note that sub-Saharan Africa's share of world exports decreased from 3.9 percent in 1980 to 1.9 percent in 2006, and the least-developed countries did even worse, with their average share falling from .06 percent to .02 percent over the period. Part of the explanation, as they point out, is that there are other nontariff barriers to trade, including supply constraints and infrastructure deficiencies, providing the rationale for "aid for trade."

15. See, in particular, Chang (2002, 2003). Moreover, developing countries that have reduced their tariffs have not been able to make up for the resulting shortage of revenues, e.g., through value-added taxes.

16. And indeed, with constraints on taxation (or subsidies), differential taxation of traded goods (as compared to domestically produced goods) is in general desirable (Dasgupta and Stiglitz 1971, 1974, 2000; Emran and Stiglitz 2005). These results are consistent with those that show that certain types of liberalization (e.g., of intermediate goods and capital goods) may have beneficial effects (see Estevadeordal and Taylor 2008). The effects of liberalization may depend too on the economy's situation and structure: when there already is a high level of unemployment, liberalization may have adverse effects, even if it has more positive effects in other circumstances (see Charlton and Stiglitz 2005).

17. Irwin and Kroszner (1999) outline the conversion of the Republican Party away from its long-time support for industrialization behind high tariff walls, beginning in the 1940s.

18. Of course, trade interventions have sometimes not worked out well. (They have been used as protectionist tools by special interests, rather than to redirect society's resources toward creating a learning society.) But the history of successful interventions suggests that failure is not inevitable. And hopefully, countries will learn from the failures (and successes) of the past, so that the returns from future interventions will presumably be greater than those from past interventions.

19. In chapter 2, we noted similarly that the existence of large unexploited potential productivity was confirmed by special historical circumstances where there was a sudden need to increase output.

20. Moreover, the circumstances confronting Latin America in the 1960s and East Asia in the 1980s and 1990s were markedly different. It is not obvious that an export growth strategy would have worked in the 1960s.

21. See Rodrik and Subramanian (2005) for the case of India. Rodrik (2001) shows that growth relative to all developing countries actually increased from 1975 to the 1980s, even though import duties increased.

22. As we noted earlier, U.S. public investments in research have had enormously high returns (Stiglitz and Wallsten 1999; Council of Economic Advisers 1995).

23. The returns on U.S. government investments in technology and science are even higher than those of the private sector (which in turn are far higher than private sector returns elsewhere; see Council of Economic Advisers 1995).

10. Financial Policy and Creating a Learning Society

1. The existence of these externalities provides the rationale for financial sector regulation, and the failure to adequately take into account these externalities provides an important part of the explanation of the 2008 and other financial crises (see Stiglitz 2010b).

2. At the center of lending activities are issues of information: assessing credit worthiness and monitoring fund usage. Markets characterized by imperfect and asymmetric information—features which are central to financial markets—are inherently characterized by externalities, resulting in market allocations not being (constrained) Pareto efficient (Greenwald and Stiglitz 1986). There are, in fact, a number of distinct categories of externalities, besides those associated with the macro-instability upon which the next chapter focuses. Actions (investments) affect credit constraints, self-selection constraints, incentive compatibility constraints, and price distributions.

3. By the Riegle-Neal Interstate Banking and Branching Efficiency Act of 1994.

4. See also Beck, Demirgüç-Kunt, and Martinez Peria (2010). We note that there are some empirical studies that claim the opposite (see Clarke, Cull, Martinez Peria, and Sanchez 2005).

5. Because of differential information, there is likely to be more subjective risk associated with a project in the developing country than in the home country, so the expected return required to induce a loan will have to be correspondingly greater. Matters may be even worse: the foreign lender may know that in competing for domestic borrowers, it faces a winner's curse. If there are local lenders with better information, the foreign lender only succeeds in "winning" if it offers a loan at too low of an interest rate—at an interest rate below the rate at which the (better) informed domestic lender is willing to lend.

6. Even borrowers will care about their lender's life expectancy. Lending is informationally intensive; borrowers develop a relationship with the lender, which makes the market for loans particularly imperfect. If a lender goes into bankruptcy in a downturn, borrowers are especially likely to find it difficult to find an alternative source of funds (see Jaffee and Stiglitz 1990; Greenwald and Stiglitz 2003).

7. The fact is that backing any country's banking system is its government; credit default swap spreads for banks and for the sovereigns of those banks are highly correlated (IMF 2012a). Argentina and the events of the global financial crisis of 2008 showed that depositor beliefs are not always fully rational: governments did not always come to the rescue in the way hoped.

8. Iceland provides an interesting case, because depositors in the U.K. and the Netherlands evidently felt that their assets were secure, though any "rational" analysis would have made clear the severe limitations in this small country's ability to protect them. The governments of the Netherlands and the U.K. put enormous pressure on Iceland, but in the end, largely failed.

9. A result that is particular startling, given the destruction of the war and the impediments posed to trade.

10. While the United States did not fully create "national banking" until the 1990s, the country's national banking system was created in 1863, with the

National Currency Act, which created a system of regulation for nationally chartered banks (the Office of the Comptroller of the Currency, under the U.S. Department of Treasury). These newly established nationally chartered banks were able to attract funds from outside the state, and though funds didn't flow as freely as they might have with banks that could operate freely across state boundaries, funds flowed more freely than in the previous regime.

11. More recently, some foreign banks have engaged in extensive consumer lending, taking advantage of their "learning" about how to better exploit uninformed consumers, replacing in some cases even more exploitative local money lenders.

12. Thus, *financial market liberalization* refers to opening up a country's markets to foreign financial institutions and the deregulation of financial markets more generally, while *capital market liberalization* focuses on the movement of capital itself into and out of a country. Discussions of capital market liberalization usually focus on short-term capital flows (bank lending, portfolio investments) rather than foreign direct investment.

13. See, e.g., Prasad et al. (2003) and Kose et al. (2006). For these authors, the fact that volatility did not decrease in many of the countries which became more financially integrated into the global economy was a puzzle. As Stiglitz (2008a) pointed out, in models with imperfect and asymmetric information (endogenous capital market imperfections), pro-cyclical capital flows could easily be explained. More generally, in models with finite-lived individuals, capital market liberalization could lead to more volatility of consumption. (The standard models assumed infinitely lived individuals.)

14. See, e.g., Stiglitz (2000a, 2002a, 2006a, 2008a); Stiglitz et al. (2006, 2008); and the references cited there.

15. Thailand provides an example. There is an important distinction between short-run flows, the major effect of which may be an increase in the exchange rate, thus discouraging export sectors, and foreign direct investment, which *may* go into sectors associated with more learning and learning externalities.

16. Earlier discussions also noted other impediments to financing research. The borrower had to disclose enough information to make the lender willing to provide money, but then the lender could "steal" the idea, appropriating the returns for himself.

17. This section borrows heavily from Greenwald and Stiglitz (2003); Hellman, Murdock, and Stiglitz (1996, 1997, 1998, 2000, 2002); Honohan and Stiglitz (2001); Murdock and Stiglitz (1993); and Stiglitz and Uy (1996).

18. Part of the reason is that information (like knowledge more generally) is a public good. If capital markets were really informationally efficient, as its advocates claim, there would be no incentive to gather information. Everyone would try to be a free rider on the investments in information of others. While

securitization may lead to improved risk diversification, it had adverse effects on incentives for assessing credit worthiness and monitoring, and these effects played out disastrously. For the general theory, see Grossman and Stiglitz (1976, 1980). For a discussion of the problems posed by securitization, and an explanation why the contention that it improves risk diversification may be incorrect, see Stiglitz (1992, 2010b).

19. Financial restraint needs to be distinguished from financial repression, which typically entailed large negative real interest rates. One of the standard arguments against financial restraint was that the lower (real) interest rates associated with it led to less savings; but interestingly, the East Asian countries all had very high savings rates. This may be partly because the interest elasticity of savings may be very low, but it also may be partly because government policies enhanced both the safety and convenience of financial savings.

20. The lower lending rates in turn helped increase firm equity, enabling them to engage in more risky investment (see Greenwald and Stiglitz 1993).

21. Moreover, it has also been more widely recognized that private banks also engage in "connected" lending, and, especially when the private bank looms large in the economy, taxpayers wind up picking up the tab. The distinction, in this sense, between public and private institutions is somewhat blurred. Monitoring public institutions may, in fact, be easier than monitoring private institutions.

22. We are suggesting, in other words, the creation of Arrow-Debreu securities related to the macro-state of the economy. Though it should be relatively easy to create these risk products, neither government nor the private sector has done so.

23. Moreover, there are limits to the interest rate that banks can charge in the initial period, because of adverse selection and incentive effects, described by Stiglitz and Weiss (1981).

24. As the prime minister of one developing country argued, this was the most important example of a taking by one country of another country's intellectual property (see Stiglitz 2006a).

11. Macroeconomic and Investment Policies for a Learning Society

1. While, typically, foreign investors are more sensitive to adverse signals, domestic investors have access to inside information. Thus, there will be some cases where that inside information indicates to them to leave before there is an adverse public signal to which foreign investors react. In the Tequila crisis of 1994/1995, it appears that Mexicans attempted to take their money out of the country first (see Lederman et al. 2003).

2. There are, of course, other explanations. For example, they have greater incentives to become tax havens, because the loss in domestic revenues from increased tax avoidance is more than offset by the increase in foreign revenues.

3. This is because weaknesses in the local banking system would not be highly correlated with weaknesses in the large international banks.

4. An exception occurs if banks become so undercapitalized that they "gamble on resurrection."

5. As we noted earlier in this chapter and in chapter 4, learning benefits from having a stable environment.

6. This is the case in most developing countries. Some critics have suggested that a low exchange rate exposes a country to more inflation. Two responses are in order: First, that would only be the case if the central bank did not take offsetting actions. When the economy is already at full employment, the exchange rate affects the composition of output, and it may still be the case that it shifts it toward the learning sectors. Second, inflation is affected not by the level of the exchange rate (which affects relative prices), but by changes in the exchange rate.

7. For a broader discussion of these issues, see, Erten, and Ocampo (2013), Gallagher (2014), Guzman, Ocampo, and Stiglitz (2014), and Ocampo and Erten (2014).

8. See Guzman et al. (2014).

9. Some investment treaties provide foreign investors greater protections than they do domestic firms (see, e.g., Stiglitz 2006a, 2008e).

10. These were among the center messages of the 1999 World Development Report, *Knowledge for Development*.

11. Moreover, individuals who are absorbed with ensuring their basic survival have less ability to learn. Good systems of social protection thus enhance individuals' learning capacities.

12. Intellectual Property

1. For a broader discussion of this issue, including empirical evidence, see Dosi, Marengo, and Pasquali (2006); and Dosi and Stiglitz (2014).

2. We should reiterate our cautionary note: more competition does not necessarily lead to more innovation. As we noted in chapters 5 and 6, however, especially because of agency problems (managerial capitalism), a monopoly may have little to spur it into innovation. This effect may dominate in markets with only one or two firms.

3. Moreover, as we noted earlier, monopoly innovators fail to take into account any consumer surplus that results from large innovations or the consumer surplus that accrues from higher levels of innovation.

4. With perfect information, presumably the owner of intellectual property could act as a perfectly discriminating monopolist and extract from potential users the surplus associated with the use of knowledge—so that there would be no distortion. But information is imperfect, and owners of intellectual property are far from perfectly discriminating monopolists. (For a discussion of imperfect information and monopoly distortion, see Stiglitz [1977].)

5. For a more extensive discussion of this point and related issues, see Mowery et al. (2001); David (2004a, 2004b); and Dosi and Stiglitz (2014).

6. He in fact used his patent to try to organize an automobile cartel. Had the patent not been challenged by Henry Ford, who wanted to create a low-priced car, the development of the automobile would have been greatly impeded. For a discussion of this and other problems with the patent system, see Stiglitz (2006a).

7. For an early discussion of the importance of the scope of the patent, see Merges and Nelson (1994).

8. There is now a large literature on this subject. See, e.g., Farrell and Shapiro (2008); Lemley and Shapiro (2007); and Shapiro (2001, 2010).

9. For a discussion of this story, see, e.g., Crouch (1989). Fighting their patent claims may also have diverted the attention of the Wright brothers from making further development in their own design, contributing to the United States falling behind Europe in the development of the airplane. The irony is that it appears that the critical insight concerning the control of the airplane had been patented decades earlier, in 1868, by British inventor Matthew Piers Watt Boulton. Had the patent examiners known of this earlier patent, they might not have granted the Wright brothers their patent. The limitations of the U.S. patent system may be further highlighted by the fact that the Wright brothers' original patent application in 1903 was rejected. It was only when they reapplied, using a patent attorney, that they were granted the patent.

10. Michael Heller and his coauthors have provided other examples of how patents can deter innovation, invoking the term *anticommons*. See Heller (1998, 2008); and Heller and Eisenberg (1998).

11. It used to be the case that once granted a patent, the owner could exclude others from using that intellectual property *until the patent was overturned*. This became a source of special concern, given the large number of bad patents—patents which should not have been issued, some of which are eventually overturned. Those who have such patents could impose extortionary demands on those who wish to make use of their patents. These patent owners can even insist that those to whom they grant license not sue—eliminating a major source of challenge to patents.

12. *eBay Inc. v. MercExchange, L.L.C.*, 547 U.S. 388 (2006). Case documents can be found at http://www.supremecourt.gov/opinions/05pdf/05–130.pdf (accessed January 15, 2013).

13. The court decision went some way to creating what intellectual property lawyers like Reichmann had long called for, a "liability system," under which those who use another's intellectual property have to pay compensation, but the owner of the intellectual property cannot exclude someone from using the property.

14. In 2012, a small company named X2Y sued Intel, Apple, and HP to exclude from the American market all of Intel's advanced microprocessors, all of Apple's computers (which employ these microprocessors), and those HP computers that do so. The claim was that these microprocessors infringed, in their "packaging," on an X2Y patent. X2Y had offered to sell this and a bundle of other patents for a few million dollars. Intel viewed it as a holdup and refused. The cost to Intel, Apple, and HP—let alone to the U.S. economy—of the exclusion would have been the order of billions of dollars. The law providing for the exclusion had a narrow exception—the exclusion order was not to be issued if it was against the public interest. But the International Trade Court (ITC) had so narrowly defined the exception that it had been used only four times in forty years. The irony, of course, was that a law designed to protect American firms against foreign firms who violated the intellectual property rights of Americans was being used by a small American firm that had spent a miniscule amount on research—and far more on lawyers—to hold up some of America's leading IT companies, who were spending billions on research. Those who argued against the exclusion order contended not only that exclusion would have a large negative effect on the economy in the short run, but also that it would be counterproductive, disincentivizing research.

15. In practice, there is usually some value to a me-too innovation—for instance, there may be some patients for whom the side-effects are less—but still, the social return to such innovations is very limited and less than the private returns.

16. See chapter 6 for a formal model demonstrating this.

17. There is a large literature on the subject. For a review, see, e.g., Gallini (2002).

18. For instance, in the case of "orphan drugs," the life of the patent was extended, because it was thought that the benefits from greater incentives to innovate exceeded the costs. A still better way of creating incentives for such innovation, however, could have been provided through the prize system.

19. In April 2010, the U.S. District Court for the Southern District of New York invalidated patents on a pair of genes linked to breast and ovarian cancer held by Myriad. But in July 2011, the Court of Appeals for the Federal Circuit overturned this decision (Pollack 2011). In 2013, the Supreme Court supported the District Court decision that one could not obtain a patent for isolating a naturally occurring gene. *Association for Molecular Pathology v. Myriad Genetics*, 569 U.S. 12–398 (2013).

20. See Brunswick, 36 Fed. Cl. at 207; cited in Love (2004, 13).

21. Differences in politics—including the influence of the pharmaceutical and entertainment industries—may, however, be the predominant explanation of the differences.

22. *United States v. W. Elec. Co.*, 1956 Trade Cas. (CCH) ¶ 68,246, at 71,139 (D.N.J. 1956). Earlier, we noted how the airplane patent pool helped resolve conflicting claims and allowed further development of the airplane.

23. Under a U.S. law called the Tunney Act (Antitrust Procedures and Penalties Act, 15 U.S.C. §16), members of the public have an opportunity to comment on a proposed settlement of a civil antitrust suit before it is accepted by a court. At the time of the proposed Microsoft settlement, Stiglitz filed an affidavit together with Jason Furman (later the chairman of the Council of Economic Advisers under President Obama) explaining why limiting the length of the patent would be a preferable way to address the anticompetitive abuses.

24. The discussion of this section is adapted from Stiglitz (2008b); Stiglitz (2013a); and Dosi and Stiglitz (forthcoming). It draws heavily upon Freeman (1987); Lundvall (2010); and Nelson (2004).

25. The general theory of prizes is set forth in Nalebuff and Stiglitz (1983a, 1983b). Subsequently, there has developed a large literature on the use of prizes as an incentive system (including Love and Hubbard 2007; Davis and Davis 2004; and the papers cited there). Also, a bill has been introduced into the U.S. Senate to use prizes as a way of incentivizing medical research. And in 2012 the WHO Consultative Expert Working Group on Research and Development: Financing and Coordination (CEWG), linked with its Commission on Intellectual Property Rights, Innovation and Public Health (set up by WHO in 2003 to look at the relationship between intellectual property, innovation, and public health), recommended establishing a prize system, as well as other measures to make medical research more "open." The CEWG also recommended creating patent pools and putting research outputs that address the health needs of developing countries into the public domain or making them available through open licensing.

26. For a recent discussion, see Kremer and Williams (2010). For an earlier discussion, see Stiglitz (2006a) and the references cited in earlier footnotes.

27. The Royal Agricultural Society in England also provided prizes. Brunt, Lerner, and Nicholas (2011) show that these prizes provided effective inducements to research.

28. Some of the benefits from using competitive markets to disseminate the knowledge can be obtained if the government buys out patents, i.e., giving the patent holder what the monopoly profits would have been (see Kremer 1998).

29. This chapter is devoted to IPR and its impact on creating a learning society. There are also questions associated with producing knowledge, e.g., whether knowledge production is best carried on in public, private for-profit,

or nonprofit institutions. The issues of production and finance largely can be separated. Production can be undertaken privately or publicly; finance can be undertaken privately or publicly. At one extreme are government research laboratories—publicly financed research that is also publicly "produced." The IPR system is often described as the polar opposite, a private-sector solution combining private funding and private finance. But this description is misleading in two respects that we have already noted: First, much of the innovation is based on basic research that is publicly funded and often publicly produced or at least produced by not-for-profit entities, such as universities. And second, in the case of both health and defense, even the seemingly "private" funding under an IPR regime is really public funding, since all defense expenditures are from the public purse and since the government provides most of the funding for health care expenditures in most countries. Even in the most market-oriented country, the United States, much of the funding comes from government: The National Institutes of Health represent publicly financed and publicly produced research; and government spending on health care, both through its program for poor people, Medicaid, and through its program for elderly people, Medicare, represents a large share of total health care spending.

30. We say "excessive" because it may in fact be optimal to have several independent, parallel research efforts.

31. The evidence is that capital markets do not fully spread risks faced by firms, because of imperfections of information. See, e.g., Greenwald and Stiglitz (1990), who discuss the effect of information imperfections on firm behavior and argue that informational problems in the capital market cause firms to act in a risk-averse manner. (See also Stiglitz 1982c). There is also considerable empirical evidence that firms act in a risk-averse manner, even when risks are uncorrelated with the market (see, e.g., Stiglitz 1982b).

32. For broader discussions of these issues, see Cimoli et al. (2014); Lewis and Reichman (2005); Nelson (2004); and Odagiri et al. (2010).

33. Statement by Brazil on September 30, 2004, before the WIPO General Assembly at the introduction of the proposal for a development agenda.

34. See, e.g., Stiglitz (2006a). Indeed, it was not even clear that the IPR regime that was foisted on the world through TRIPS was well designed for the United States, as we suggested in the beginning of the chapter. It reflected the interests of the entertainment and pharmaceutical industries, not of the scientific community.

35. For instance, in granting pharmaceutical patents, developing countries should reserve the right to grant a compulsory license for any lifesaving or life-extending drug. To be exempt from this provision, the patent applicant would have to state that the patent does not cover any such medicinal use; and if subsequently such a use were established, the government would have the right to issue a compulsory license, limited, of course, to sales for such usages.

In the context of trade agreements, see Charlton and Stiglitz (2012) and Ismail (2007) for a discussion of the "right to development."

36. See also Odagiri et al. (2010) and the various chapters of Cimoli et al. (2014).

13. Social Transformation and the Creation of a Learning Society

1. As we also noted in chapter 3, in some quarters and in some countries, it appears that the notion that policies ought to be based on the principles of the Enlightenment has to be constantly relitigated.

2. For instance, a classic experiment in psychology by Bruner and Potter (1964) suggests that preconceived ideas serve as unconscious filters of sensory impressions.

3. *Confirmatory bias* is the tendency to search for, interpret, and remember information in a way that supports one's initial beliefs. For a survey, see Rabin and Schrag (1999).

4. On the other hand, it provides a more disciplined approach to the formation of beliefs than that based on "animal spirits," which suggests that any set of beliefs is possible.

5. In this sense, our analysis goes beyond standard behavioral economics, which has used insights from psychology to modify economists' traditional reliance on hypotheses concerning individual rationality.

6. Of course, sociologists have long recognized the importance of social constructions (see, e.g., Douglas 1986), but they have not focused on modeling "equilibrium," where there is some correspondence between beliefs and perceptions and what the individuals observe.

Some economic historians have also emphasized ideas similar to those articulated here (see, e.g., North 2005).

7. Standard theory treats the categories as if they were objectively determined. Standard rational expectations theory assumes that individuals use all the relevant information, updating prior beliefs through a Bayesian process. There are no biases.

This approach is also markedly different from the very interesting models of Piketty (1995); Bénabou (2008a, 2008b); and Bénabou and Tirole (2002), who assume that individuals *strategically* choose the probability that they will remember certain signals.

8. There are an infinite number of possible correlations between observables. Individuals have to choose which among these they study. They do not gather information about many of these possible correlations because the way we see the world suggests that they are irrelevant. If we came to believe that they were relevant, they possibly would be. This is called preconfirmatory bias.

Fryer and Jackson (2008) analyze bias that emerges from categorization. See also Loury (2002).

9. Again, there is a large literature in both psychology and economics consistent with this hypothesis. Smith et al. (2008) showed that invoking in experimental subjects the feeling that they have little power impairs their performance in complex cognitive tasks. Steele (2010) provides a survey of the literature demonstrating that cuing an identity associated with a stereotype, or cuing a condition that could confirm a negative stereotype, shifts an individual's performance in the direction of the stereotype. See also Hoff and Pandey (2011); and Afridi, Li, and Ren (2011). Experiments summarized by Compte and Postlewaite (2004) demonstrate that psychological states can affect performance. Among the earliest examples are the efficiency wage theories in economics, which noted that perceptions of unfairness can affect morale, which can in turn affect performance (see Stiglitz 1974b; and Akerlof and Yellen 1986).

10. Efficiency wage theory (referred to in an earlier footnote) provided early examples of this. Perceptions of fairness can affect morale, morale can affect behavior, and this can explain the persistence of dysfunctional inequality.

11. The possibility of multiple equilibria of this sort is enhanced once it is recognized that a more dynamic society enhances the returns associated with innovative skills and attributes, while in a less dynamic society the relative returns of bureaucrats may be higher. A high-learning society creates an ecology which is self-supporting.

12. See in particular the report of the National Commission on the Causes of the Financial and Economic Crisis in the United States (2011).

13. The complexity of the issues is illustrated by vicissitudes in attitudes toward government policies to restimulate the economy. In the aftermath of the collapse of Lehman Brothers, there was a moment in which all the world adhered to Keynesian ideas. But within two years, there was a shift toward "Hooverite" fiscal austerity policies—even though the empirical (scientific) evidence that such policies would lead to slower growth with disappointing results on deficit reduction had actually mounted in the interim.

14. But it is not as if those who believe in that institution—or even the smaller group that benefits from it—got together and figured out a set of beliefs that would accomplish what they sought. As we have noted, the theory we have presented is incomplete, in that it does not adequately explain when beliefs change and when they do not. But we believe it is a step forward to break out of the mold of rational expectations, in which the variables described earlier play no role.

15. Even the way we perceive institutions is affected by the prisms through which we look at the world, by our ideology. At one time, some economists suggested that institutions have a simple role in society—to fill in the "holes" in markets, to remedy market failures (see North 1971). Arnott and Stiglitz

(1991); and Hoff and Sen (2006), as well as others, showed that nonmarket institutions purportedly resolving a market failure (like incomplete insurance markets) could, in this sense, be dysfunctional—they could lead to Pareto inferior outcomes.

More recent literature has highlighted the role of institutions in preserving inequalities—in the context of repeated games, equilibria in which one group is exploited by others may be sustained (see, among others, Dasgupta 2005; and Mookherjee and Ray 2003).

16. We are deeply indebted to Tim Besley for discussions on the ideas in this section (see Besley and Persson 2009, 2010, forthcoming; and Besley, Persson, and Sturm 2010). Hoff and Stiglitz (2004a, 2004b, 2007) modeled the political economy of transition from communism to a market economy, employing analogous ideas. See also Acemoglu and Robinson (2000).

14. Concluding Remarks

1. See Stiglitz (2011) and the references cited there.

2. As we noted earlier, there can be some knife-edge circumstances, where firms begin in precisely identical situations, and remain so. But these are highly unstable. Any perturbation that leads one firm to have any advantage over others will have cumulative effects, until that firm becomes dominant.

3. We emphasized in earlier chapters that the market is *constrained* Pareto inefficient—even taking into account the costs of gathering information or removing information asymmetries.

Index

References

Abernathy, William J., and Kim B. Clark. 1985. "Innovation: Mapping the Winds of Creative Destruction." *Research Policy* 14 (1): 3–22.

Acemoglu, Daron. 2010. "When Does Labor Scarcity Encourage Innovation?" *Journal of Political Economy* 118 (6): 1037–1078.

Acemoglu, Daron, and James Robinson. 2012. *Why Nations Fail*. New York: Random House.

——. 2000. "Why Did the West Extend the Franchise? Democracy, Inequality and Growth in Historical Perspective." *Quarterly Journal of Economics* 115 (4): 1167–1199.

Acemoglu, Daron, Philippe Aghion, L. Bursztyn, and D. Hemous. 2009. "The Environment and Directed Technical Change." NBER Working Paper 15451. http://www.nber.org/papers/w15451.

Acemoglu, Daron, Philippe Aghion, and Fabrizio Zilibotti. 2006. "Distance to Frontier, Selection and Economic Growth." *Journal of the European Economic Association* 4 (1): 37–74.

Acemoglu, Daron, James Robinson, and Thierry Verdier. 2012. "Can't We All Be More Like Scandinavians? Asymmetric Growth and Institutions in an Interdependent World." NBER Working Paper 18441. http://www.nber.org/papers/w18441.pdf.

Africa Progress Panel. 2012. *Africa Progress Report: Jobs, Justice and Equity*. Available online: http://www.africaprogresspanel.org/publications/policy-papers/africa-progress-report-2012/.

Afridi, Farzana, Sherry Xin Li, and Yufei Ren. 2011. "Social Identity and Inequality: The Impact of China's Hukou System." Manuscript, University of Texas at Dallas. http://ideas.repec.org/p/cde/cdewps/190.html.

Aghion, Philippe. 1999. "Development Banking." *Journal of Development Economics* 58: 83–100.

Aghion, Philippe, and Peter Howitt. 1998. *Endogenous Growth Theory*. Cambridge: Massachusetts Institute of Technology Press.

———. 1992. "A Model of Growth through Creative Destruction." *Econometrica* 60 (2): 323–351.

Aghion, Philippe, Peter Howitt, and Susanne Prantl. 2013. "Patent Rights, Product Market Reforms, and Innovation." Working Paper, Harvard University.

Aghion, Philippe, Ufuk Akcigit, and Peter Howitt. 2013. "What Do We Learn from Schumpeterian Growth Theory?" NBER Working Paper 18824.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt. 2005. "Competition and Innovation: An Inverted U Relationship." *Quarterly Journal of Economics* 120 (2): 701–728.

Aghion, Philippe, Robin Burgess, Stephen Redding, and Fabrizio Zilibotti. 2008. "The Unequal Effects of Liberalization: Evidence from Dismantling the License Raj in India." *American Economic Review* 98 (4): 1397–1412.

Aghion, Philippe, Mathias Dewatripont, L. Du, A. Harrison, and P. Legros. 2012. Industrial Policy and Competition. Manuscript, Harvard University. http://scholar.harvard.edu/files/aghion/files/industrial_policy_and_competition.pdf.

Aghion, Philippe, David Hemous, and Enisse Kharroubi. 2009. "Cyclical Budgetary Policy, Credit Constraints, and Industry Growth." Working Paper, Harvard University. http://scholar.harvard.edu/aghion/publications/credit-constraints-cyclical-fiscal-policy-and-industry-growth.

Ahmad, Syed. 1966. "On the Theory of Induced Invention." *Economic Journal* 76: 344–357.

Aizenman, Joshua, and Yothin Jinjarak. 2009. "Globalisation and Developing Countries—a Shrinking Tax Base?" *Journal of Development Studies* 45 (5): 653–671.

Akerlof, George. 1970. "The Market for 'Lemons': Qualitative Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 89: 488–500.

Akerlof, George A., and R. E. Kranton. 2010. *Identity Economics: How Our Identity Shapes Our Work, Wages and Well-Being*. Princeton, N.J.: Princeton University Press.

Akerlof, George A., and Joseph E. Stiglitz. 1969. "Capital, Wages and Structural Unemployment." *Economic Journal* 79 (314): 269–281.

Akerlof, George A., and Janet L. Yellen. 1990. "The Fair Wage-Effort Hypothesis and Unemployment." *Quarterly Journal of Economics* 105 (2): 255–283.

———. 1986. *Efficiency Wage Models of the Labor Market*. Cambridge: Cambridge University Press.

Alchian, Armen. 1963. "Reliability of Progress Curves in Airframe Production." *Econometrica* 31: 679–693.

Alloy, Lauren B., and Lyn Y. Abramson. 1979. "Judgment of Contingency in Depressed and Nondepressed Students: Sadder but Wiser?" *Journal of Experimental Psychology: General* 108 (4): 441–485.

Aoki, Masahiko. 1970. "A Note on Marshallian Process under Increasing Returns." *Quarterly Journal of Economics* 84 (1): 100–112.

Argote, Linda, and Dennis Epple. 1990. "Learning Curves in Manufacturing." *Science* 247: 920–924.

Argote, Linda, Sara L. Beckman, Dennis Epple. 1990. "The Persistence and Transfer of Learning in Industrial Settings." *Management Science* 36 (2): 140–154.

Ariely, Dan. 2008. *Predictably Irrational: The Hidden Forces That Shape Our Decision*s. Rev. and expanded ed. New York: Harper Collins.

Arnott, Richard, and Joseph E. Stiglitz. 1991. "Moral Hazard and Nonmarket Institutions: Dysfunctional Crowding Out or Peer Monitoring?" *American Economic Review* 81 (1): 179–190.

——. 1985. "Labor Turnover, Wage Structure & Moral Hazard: The Inefficiency of Competitive Markets." *Journal of Labor Economics* 3 (4): 434–462. Reprinted in *Selected Works of Joseph E. Stiglitz*, vol. 2, *Information and Economic Analysis: Applications to Capital, Labor, and Product Markets*, 581–600. Oxford: Oxford University Press, 2013.

Arnott, Richard, Bruce Greenwald, and Joseph E. Stiglitz. 1994. "Information and Economic Efficiency." *Information Economics and Policy* 6 (1): 77–88.

Arnott, Richard, A. Hosios, and Joseph E. Stiglitz. 1988. "Implicit Contracts, Labor Mobility and Unemployment." *American Economic Review* 78 (5): 1046–1066.

Arrow, Kenneth J. 1965. *Aspects on the Theory of Risk-Bearing*. Helsinki: Yrjö Jahnssonin Säätiö.

——. 1962a. "The Economic Implications of Learning by Doing." *Review of Economic Studies* 29: 155–173.

——. 1962b. "Economic Welfare and the Allocation of Resources for Invention." In *The Rate and Direction of Inventive Activity: Economic and Social Factors*, ed. R. Nelson, National Bureau of Economic Research (NBER): 609–629. Princeton, N.J.: Princeton University Press.

——. 1951a. "An Extension of the Basic Theorems of Classical Welfare Economics." In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, ed. J. Neyman, 507–532. Berkeley: University of California Press.

——. 1951b. *Social Choice and Individual Values*. New York: Wiley.

Arrow, Kenneth J., and F. Debreu. 1954. "Existence of an Equilibrium for a Competitive Economy." *Econometrica* 22: 265–290.

Arrow, Kenneth J., W. R. Cline, K-G. Maler, M. Munasinghe, R. Squitieri, and Joseph E. Stiglitz. 1996. "Intertemporal Equity, Discounting, and Economic Efficiency." In *Climate Change 1995: Economic and Social Dimensions of Climate Change*, ed. J. Bruce, H. Lee, and E. Haites, 21–51. Cambridge: Cambridge University Press.

Arvis, Jean-François. 2013. "How Many Dimensions Do We Trade In? Products Space Geometry and Latent Comparative Advantage." World Bank Policy Research Working Paper 6478. http://elibrary.worldbank.org/doi/pdf/10.1596/1813-9450-6478.

Asher, Harold. 1956. Cost-Quantity Relationships in the Airframe Industry. Paper R-291, RAND Corp., Santa Monica.

Atkinson, Anthony B., and Joseph E. Stiglitz. 1980. *Lectures on Public Economics* New York: McGraw-Hill.

——. 1969. "A New View of Technological Change." *Economic Journal* 79 (315): 573–578.

Autor, David H., and David Dorn. 2013. "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market." *American Economic Review* 103 (5): 1553–1597.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics* 90 (2): 300–323.

——. 2006. "The Polarization of the U.S. Labor Market." *American Economic Review* 96 (2): 189–194.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Investigation." *Quarterly Journal of Economics* 118 (4): 1279–1333.

Azvolinsky, Anna. 2012. "Lack of BRCA Testing Approval Creates Snag for Cancer Trials." *Nature Medicine* 18 (1310).

Baily, Martin Neil, and Robert M. Solow. 2001. "International Productivity Comparisons Built from the Firm Level." *Journal of Economic Perspectives* 15 (3): 151–172.

Baily, Martin Neil, Charles Hulten, David Campbell, Timothy Bresnahan, and Richard E. Caves. 1992. "Productivity Dynamics in Manufacturing Plants." *Brookings Papers on Economic Activity: Microeconomics*: 187–267.

Baldwin, R. 1969. "The Case Against Infant-Industry Tariff Protection." *Journal of Political Economy* 77 (3): 295–305.

Barrios, Salvador, and Eric Strobl. 2004. "Learning by Doing and Spillovers: Evidence from Firm-Level Panel Data." *Review of Industrial Organization* 25 (2): 175–203.

Barro, R.J. & Sala-I-Martin, X., 1991. "Convergence Across States and Regions," Papers 629, Yale Economic Growth Center.

Barth, Erling, B. Finseraas, and K. O. Moene. 2012. "Political Reinforcement: How Rising Inequality Curbs Manifested Welfare Generosity." University of Oslo.

Barth, Erling, Karl O. Moene, and Fredrik Willumsen. 2014. "The Scandinavian Model: An Interpretation." *Journal of Public Economics*, 117: 60–72.

Baumol, William J. 1982. "Contestable Markets: An Uprising in the Theory of Industry Structure." *American Economic Review* 72 (March): 1–15.

Baumol, William J., John C. Panzar, and Robert D. Willig. 1982. *Contestable Markets and the Theory of Industry Structure*. New York: Harcourt, Brace and Jovanovich.

Bayraktar, Nihal, and Yan Wang. 2004. "Foreign Bank Entry, Performance of Domestic Banks and Sequence of Financial Liberalization." Policy Research Working Paper Series 3416, Washington, D.C.: World Bank.

Beck, Thorsten, and Asli Demirgüç-Kunt. 2006. "Small and Medium-Size Enterprises: Access to Finance as a Growth Constraint." *Journal of Banking and Finance* 30: 2931–2943.

Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine. 2000. "A New Database on Financial Development and Structure." *World Bank Economic Review* 14: 597–605.

Beck, Thorsten, Asli Demirgüç-Kunt, and Vojislav Maksimovic. 2008. "Financing Patterns Around the World: Are Small Firms Different?" *Journal of Financial Economics* 89: 467–487.

Beck, Thorsten, Asli Demirgüç-Kunt, and Maria Soledad Martinez Peria. 2010. "Foreign Banks and Small and Medium Enterprises: Are They Really Estranged?" http://www.voxeu.org/index.php?q=node/4828.

Bénabou, Roland. 2008a. Groupthink: Collective Delusions in Organizations and Markets. Manuscript, Princeton University, Princeton, N.J.

——. 2008b. "Ideology." *Journal of the European Economic Association* 6 (2–3): 321–352.

——. 1996. "Inequality and Growth." In *NBER Macroeconomics Annual 1996*, vol 11, ed. Ben S. Bernanke and Julio J. Rotemberg, 11–92. Cambridge, Mass.: MIT Press.

Bénabou, Roland, and Jean Tirole. 2009. "Over My Dead Body: Bargaining and the Price of Dignity." In *Papers and Proceedings of the One Hundred Twenty-First Meeting of the American Economic Association*, *American Economic Review* 99 (2): 459–465.

——. 2006. "Incentives and Prosocial Behavior." *American Economic Review* 96 (5): 1652–1678.

——. 2003. "Intrinsic and Extrinsic Motivation." *Review of Economic Studies* 70 (3): 489–520.

——. 2002. "Self-Confidence and Personal Motivation." *Quarterly Journal of Economics* 117: 871–915.

Berg, Andrew, and Jonathan Ostry. 2011. "Inequality and Unsustainable Growth: Two Sides of the Same Coin?" IMF Staff Discussion Note No. 11/08, April, International Monetary Fund.

Bértola, Luis, and José Antonio Ocampo. 2012. "Learning from Latin America: Debt Crises, Debt Rescues and When and Why They Work." Institute for the Studies of the Americas Working Paper, Institute for the Study of the Americas.http://americas.sas.ac.uk/sites/default/files/files/filestore-documents/events/Papers/Bertola_and_Ocampo_paper.pdf.

Besanko, David, Ulrich Doraszelski, Yaroslav Kryukov, and Mark Satterthwaite. 2010. "Learning-by-Doing, Organizational Forgetting, and Industry Dynamics." *Econometrica* 78: 453–508.

Besley, Tim, and Torsten Persson. 2011. "The Logic of Political Violence." *Quarterly Journal of Economics* 126 (3): 1411–1445.

——. 2010. "State Capacity, Conflict and Development." *Econometrica* 78 (1): 1–34.

——. 2009. "The Origins of State Capacity: Property Rights, Taxation and Politics." *American Economic Review* 99 (4): 1218–1244.

Besley, Tim, Torsten Persson, and Daniel M. Sturm. 2010. "Political Competition, Policy and Growth: Theory and Evidence from the United States." *Review of Economic Studies* 77 (4): 1329–1352.

Bessen, James, and Michael J. Meurer. 2008. *Patent Failure: How Judges, Bureaucrats, and Lawyers Put Innovators at Risk*. Princeton, N.J.: Princeton University Press.

Biais, B., and P. Woolley. 2011. "High Frequency Trading." IDEI Working Paper. http://idei.fr/doc/conf/pwri/biais_pwri_0311.pdf.

Bittlingmayer, George. 1988. "Property Rights, Progress, and the Aircraft Patent Agreement." *Journal of Law and Economics* 31 (1): 227–248.

Blitzer, C., P. Dasgupta, and Joseph E. Stiglitz. 1981. "Project Appraisal and Foreign Exchange Constraints." *Economic Journal* 91 (361): 58–74.

Boldrin, Michele, and David K. Levine. 2013. "The Case against Patents." *Journal of Economic Perspectives* 27 (1): 3–22.

Boyle, James. 2008. *The Public Domain: Enclosing the Commons of the Mind*. New Haven, Conn.: Yale University Press.

——. 2003. "The Second Enclosure Movement and the Construction of the Public Domain." *Law and Contemporary Problems* 66 (33): 33–74.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang. 2012. "Creative Accounting or Creative Destruction? Firm-Level Productivity Growth in Chinese Manufacturing." *Journal of Development Economics* 97 (2): 339–351.

Braverman, A., and Joseph E. Stiglitz. 1986. "Landlords, Tenants and Technological Innovations." *Journal of Development Economics* 23 (2): 313–332.

Bruce, James P., Hoesung Lee, and Erik F. Haites, eds. 1996. *Climate Change 1995: Economic and Social Dimensions of Climate Change*. Cambridge: Cambridge University Press.

Bruner, J., and M. C. Potter. 1964. "Interference in Visual Recognition." *Science* 144 (3617): 424–425.

Brunt, Liam, Josh Lerner, and Tom Nicholas. 2011. "Inducement Prizes and Innovation." NHH Dept. of Economics Discussion Paper No. 25/2011. http://ssrn.com/abstract=1972290; or http://dx.doi.org/10.2139/ssrn.1972290.

Cabral, Luis, and Michael Riordan. 1994. "The Learning Curve, Market Dominance, and Predatory Pricing." *Econometrica* 62: 1115–1140.

Cass, David., and Joseph E. Stiglitz. 1969. "The Implications of Alternative Saving and Expectations Hypotheses for Choices of Technique and Patterns of Growth." *Journal of Political Economy* 77 (4), Part 2: 586–627.

Chang, Ha Joon. 2003. "Kicking Away the Ladder: Infant Industry Promotion in Historical Perspective." *Oxford Development Studies* 31 (1): 21–32.

——. 2002. *Kicking Away the Ladder—Development Strategy in Historical Perspective*. London: Anthem Press.

——. 2001. "Intellectual Property Rights and Economic Development: Historical Lessons and Emerging Issues." *Journal of Human Development and Capabilities* 2 (2): 287–309.

Charlton, A. and Joseph E. Stiglitz. 2013. "The Right to Trade: Rethinking the Aid for Trade Agenda," in *Assessing Aid for Trade: Effectiveness, Current Issues and Future Directions*, ed. Mohammad A.Razzaque and Dirk Willem te Velde. London: Commonwealth Secretariat, 359–386.

Charlton, Andrew, and Joseph E. Stiglitz. 2012. "The Right to Trade: A Report for the Commonwealth Secretariat on Aid for Trade." London: Commonwealth Secretariat, August.

——. 2006. "Aid for Trade: A Report for the Commonwealth Secretariat." London: Commonwealth Secretariat. Published in *International Journal of Development Issues* 5 (2): 1–41; abridged version in *Swiss Review of International Economic Relations* 61 (2): 2006.

——. 2005. *Fair Trade for All*. New York: Oxford University Press.

Cimoli, Mario, Giovanni Dosi, Keith E. Maskus, Ruth L. Okediji, Jerome H. Reichman, and Joseph E. Stiglitz, eds. 2014. *Intellectual Property Rights and Development*. Oxford: Oxford University Press.

Cimoli, Mario, Giovanni Dosi, and Joseph E. Stiglitz, eds. 2009. *Industrial Policy and Development: The Political Economy of Capabilities Accumulation*. New York: Oxford University Press.

Clarke, George, Robert Cull, Maria Soledad Martinez Peria, and Susana M. Sanchez. 2005. "Bank Lending to Small Businesses in Latin America: Does Bank Origin Matter?" *Journal of Money, Credit and Banking* 37 (1): 83–118.

Clarkson, G., and DeKorte, D. 2006. The Problem of Patent Thickets in Convergent Technologies. *Annals of the New York Academy of Sciences*, 1093(1), 180–200.

Coase, R. 1937. "The Nature of the Firm." *Economica* 4 (16): 386–405.

Commission on Growth and Development. 2008. "The Growth Report: Strategies for Sustained Growth and Inclusive Development." The International Bank for Reconstruction and Development/The World Bank, Washington D.C. http://web.worldbank.org/WBSITE/EXTERNAL/EXTABOUTUS/ORGANIZATION/EXTPREMNET/0,,contentMDK:23225680ÐpagePK:64159605ÐpiPK:64157667ÐtheSitePK:489961,00.html.

Compte, Oliver, and Andrew Postlewaite. 2004. "Confidence-Enhanced Performance." *American Economic Review* 94 (5): 1535–1557.

Conniff, R. 2011 "King Ludd's War: In Luddite Protests, Which Began 200 Years Ago This Month, Technology Wasn't Really the Enemy." *Smithsonian Magazine* 41 (11): 82.

Corlett, W. J., and D. C. Hague. 1953. "Complementarity and the Excess Burden of Taxation," *Review of Economic Studies* 21: 21–30.

Council of Economic Advisers. 1995. "Supporting Research and Development to Promote Economic Growth: The Federal Government's Role." White Paper, October. http://clinton1.nara.gov/White_House/EOP/CEA/econ/html/econ-top.html.

Crouch, Tom D. 1989. *The Bishop's Boys: A Life of Wilbur and Orville Wright*. New York: Norton.

Dasgupta, Partha S. 2005. "Economics of Social Capital." *Economic Record* 81 (s1): S2–S21.

———. 2001. *Human Well-Being and the Natural Environment*. New York: Oxford University Press.

———. 1993. *An Inquiry into Well-Being and Destitution*. New York: Oxford University Press.

———. 1969. "On the Concept of Optimum Population." *Review of Economic Studies* 36 (107): 295–318.

Dasgupta, Partha, and Paul David. 1994. "Toward a New Economics of Science." *Research Policy* 23 (5): 487–521.

Dasgupta, Partha S., and Geoffrey M. Heal. 1979. *Economic Theory and Exhaustible Resources*. Cambridge: Cambridge University Press.

———. 1974. "The Optimal Depletion of Exhaustible Resources." *Review of Economic Studies* 41, *Symposium on the Economics of Exhaustible Resources*, 3–28. Edinburgh, Scotland: Longman Group.

Dasgupta, Partha S., and Joseph E. Stiglitz. 2000. "Formal and Informal Institutions." In *Social Capital: A Multifaceted Perspective*, ed. P. Dasgupta and I. Serageldin, 59–68. Washington, D.C.: World Bank.

———. 1988a. "Learning by Doing, Market Structure, and Industrial and Trade Policies." *Oxford Economic Papers* 40 (2): 246–268.

———. 1988b. "Potential Competition, Actual Competition and Economic Welfare." *European Economic Review* 32 (May): 569–577.

———. 1982. "Market Structure and Resource Depletion: A Contribution to the Theory of Intertemporal Monopolistic Competition." *Journal of Economic Theory* 28 (1): 128–164.

———. 1981a. "Entry, Innovation, Exit: Toward a Dynamic Theory of Oligopolistic Industrial Structure." *European Economic Review* 15 (2): 137–158.

———. 1981b. "Market Structure and Resource Extraction Under Uncertainty." *Scandinavian Economic Journal* 83: 318–333.

———. 1981c. "Resource Depletion Under Technological Uncertainty." *Econometrica* 49 (1): 85–104.
———. 1980a. "Industrial Structure and the Nature of Innovative Activity." *Economic Journal* 90 (358): 266–293.
———. 1980b. "Uncertainty, Market Structure and the Speed of R&D." *Bell Journal of Economics* 11 (1): 1–28.
———. 1977. "Tariffs Versus Quotas As Revenue Raising Devices Under Uncertainty." *American Economic Review* 67 (5): 975–981.
———. 1974. "Benefit-Cost Analysis and Trade Policies." *Journal of Political Economy* 82 (1): 1–33.
———. 1972. "On Optimal Taxation and Public Production." *Review of Economic Studies* 39 (1): 87–103.
———. 1971. "Differential Taxation, Public Goods, and Economic Efficiency." *Review of Economic Studies* 38 (2): 151–174.
Dasgupta, Partha S., Richard Gilbert, and Joseph E. Stiglitz. 1983. "Strategic Considerations in Invention and Innovation: The Case of Natural Resources." *Econometrica* 51 (5): 1430–1448.
———. 1982. "Invention and Innovation Under Alternative Market Structures: The Case of Natural Resources." *Review of Economic Studies* 49 (4): 567–582.
———. 1981. "Energy Resources and Research and Development." In *Erschopfbare Ressourcen*, ed. Horst Siebert, 85–108. Berlin: Duncker and Humbolt.
Dasgupta, Partha S., Geoffrey M. Heal, and Joseph E. Stiglitz. 1980. "The Taxation of Exhaustible Resources." In *Public Policy and the Tax System*, ed. G. A. Hughes and G. M. Heal, 150–172. London: George Allen and Unwin.
Dasgupta, Partha S., Geoffrey M. Heal, Joseph E. Stiglitz, Richard Gilbert, and David Newbery. 1977. *An Economic Analysis of the Conservation of Depletable Natural Resources*. Prepared for the Federal Energy Administration, May.
David, Paul A. 2004a. "From Keeping Nature's Secrets to the Institutionalization of Open Science." In *Collaborative Ownership and the Digital Economy (CODE)*, ed. R. A. Ghosh, 85–106. Cambridge, Mass.: MIT Press.
———. 2004b. "Understanding the Emergence of 'Open Science' Institutions: Functionalist Economics in Historical Context." *Industrial and Corporate Change* 13: 571–589.
———. 2002. "Does the New Economy Need All the Old IPR Institutions? Digital Information Goods and Access to Knowledge for Economic Development." Presented at Wider Conference on the New Economy in Development, Helsinki.
———. 1993. "Intellectual Property Institutions and the Panda's Thumb: Patents, Copyrights, and Trade Secrets in Economic Theory and History." In *Global Dimensions of Intellectual Property Rights in Science*, ed. M. B. Wallerstein, M. E. Mogee, and R. A. Schoen, 19–62. Washington, D.C.: National Academies Press.

——. 1975. *Technical Choice Innovation and Economic Growth: Essays on American and British Experience in the Nineteenth Century*. London: Cambridge University Press.

Davis, Lee, and Jerome Davis. 2004. "How Effective Are Prizes as Incentives to Innovation? Evidence from Three 20th Century Contests." Paper presented at the DRUID Summer Conference, May 7.

Debreu, G. 1959. *The Theory of Value*. New Haven, Conn.: Yale University Press.

——. 1952. "Market Equilibrium." *Proceedings of the National Academy of Sciences* 42 (1956): 876–878.

De la Fuente, Angel, and Rafael Doménech. 2006. "Human Capital in Growth Regressions: How Much Difference Does Data Quality Make?" *Journal of the European Economic Association* 4 (1): 1–36.

De Long, J. B. 1988. "Productivity Growth, Convergence, and Welfare: Comment," *The American Economic Review*, 78(5): 1138–1154.

Delli Gatti, Domenico, Mauro Gallegati, Bruce Greenwald, Alberto Russo, and Joseph E. Stiglitz. 2013. "Sectoral Imbalances and Long Run Crises." In *The Global Macro Economy and Finance*, ed. Franklin Allen, Masahiko Aoki, Jean-Paul Fitoussi, Nobuhiro Kiyotaki, Richard Gordon, and Joseph E. Stiglitz, 61–97. IEA Conference Volume No. 150-III. Houndmills, UK and New York: Palgrave.

——. 2012. "Mobility Constraints, Productivity Trends, and Extended Crises." *Journal of Economic Behavior & Organization* 83 (3): 375–393.

Denison, Edward F. 1962. *The Sources of Economic Growth in the United States and the Alternatives Before Us*. New York: Committee for Economic Development.

Detragiache, Enrica, Thierry Tressel, and Poonam Gupta. 2008. "Foreign Banks in Poor Countries: Theory and Evidence." *Journal of Finance* 63 (5): 2123–2160.

Diamond, P. 1967. "The Role of a Stock Market in a General Equilibrium Model with Technological Uncertainty." *American Economic Review* 57: 753–776.

Diamond, Peter A., and James Mirrlees. 1971a. "Optimal Taxation and Public Production I: Production Efficiency." *American Economic Review* 61: 8–27.

——. 1971b. "Optimal Taxation and Public Production II: Tax Rules." *American Economic Review* 61: 261–278.

Dick, Andrew R. 1991. "Learning by Doing and Dumping in the Semiconductor Industry." *Journal of Law and Economics* 34 (1): 133–159.

Dickey, D. A., and Fuller, W. A. 1979. "Distribution of the Estimators for Autoregressive Time Series with a Unit Root". *Journal of the American Statistical Association* 74 (366): 427–431.

Dixit, A. 2012. "Corruption: Supply-Side and Demand-Side Solutions." Text for Silver Jubilee Lecture at the Indira Gandhi Institute for Development Research, February.

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

Dollar, D. 1992. "Outward-Oriented Developing Economies Really Do Grow More Rapidly: Evidence from 95 LDCs, 1976–85." *Economic Development and Cultural Change* 40 (3): 523–544.

Domar, E. D., and R. A. Musgrave. 1944. "Proportional Income Taxation and Risk-Taking." *Quarterly Journal Economics* 58: 388–422.

Dosi, Giovanni, and Joseph E. Stiglitz. 2014. "The Role of Intellectual Property Rights in the Development Process, with Some Lessons from Developed Countries: An Introduction." In *Intellectual Property Rights and Development*, ed. M. Cimoli, G. Dosi, K. Maskus, R. Okediji, J. Reichman, and Joseph Stiglitz. Oxford: Oxford University Press.

Dosi, Giovanni, L. Marengo, and C. Pasquali. 2006. "How Much Should Society Fuel the Greed of Innovators? On the Relations Between Appropriability, Opportunities and Rates of Innovation." *Research Policy* 35 (8): 1110–1121.

Dosi, Giovanni, Richard Nelson, Christopher Freeman, Luc Soete, and Gerald Silverberg, eds. 1988. *Technical Change and Economic Theory*. London: Pinter.

Douglas, Mary. 1986. *How Institutions Think*. Syracuse, N.Y.: Syracuse University Press.

Dowrick, S. and D. T. Nguyen. 1989. "OECD Comparative Economic Growth 1950–85: Catch-Up and Convergence," *American Economic Review* 79(5): 1010–1030.

Drandakis, Emmanuel, and Edmund S. Phelps. 1966. "A Model of Induced Invention, Growth, and Distribution." *Economic Journal* 76 (December): 832–840.

Durlauf, S. N., Quah, D. T. 1999. The New Empirics of Economic Growth. In *Handbook of Macroeconomics*,vol. 1, ed. Taylor, J. B., Woodford, M. New York: Elsevier Science.

Dwyer, Douglas. 1998. "Technology Locks, Creative Destruction, and Nonconvergence in Productivity Levels." *Review of Economic Dynamics* 1 (2): 430–473.

Edgeworth, Francis Y. 1925. *Papers Relating to Political Economy*, vol. 3. London: Macmillan.

Elborgh-Woytek, Katrin, Jean-Jacques Hallaert, Hans P. Lankes, Azim Sadikov, and Dustin Smith. 2006. "Fiscal Implications of Multilateral Tariff Cuts." IMF Working Paper WP/06/203, Washington, D.C.

Ellerman, David, and Joseph E. Stiglitz. 2001. "Not Poles Apart: 'Whither Reform?' and 'Whence Reform?'" *Journal of Policy Reform* 4 (4): 325–338.

——. 2000. "New Bridges Across the Chasm: Macro- and Micro-Strategies for Russia and Other Transitional Economies." In *Zagreb International Review of Economics and Business* 3 (1): 41–72.

Emran, Shahe, and Joseph E. Stiglitz. 2009. Financial Liberalization, Financial Restraint, and Entrepreneurial Development. Institute for International Economic Policy Working Paper Series, Elliott School of International Affairs, The George Washington University, January.

———. 2005. "On Selective Indirect Tax Reform in Developing Countries." *Journal of Public Economics*, April, 599–623.
Erten, Bilge, and José Antonio Ocampo. 2013. "Capital Account Regulations, Foreign Exchange Pressure, and Crisis Resilience." Initiative Policy Dialogue Working Paper.
Estevadeordal, Antoni, and Alan Taylor. 2008. Is the Washington Consensus Dead? Growth, Openness, and the Great Liberalization, 1970s–2000s. NBER Working Paper 14264.
European Commission. 2008. "Pharmaceutical Sector Inquiry: Preliminary Report." DG Competition Staff Working Paper 9.
Farrell, Joseph. 1987. "Information and the Coase Theorem." *Journal of Economic Perspectives* 1: 113–129.
———. 1986. "How Effective Is Potential Competition?" *Economics Letters* 20: 67–70.
Farrell, Joseph, and Carl Shapiro. 2008. "How Strong Are Weak Patents?" *American Economic Review* 98 (4): 1347–1369.
———. 1988. "Dynamic Competition with Switching Costs." *RAND Journal of Economics* 19 (1): 123–137.
Fellner, William. 1961. "Two Propositions in the Theory of Induced Innovations." *Economic Journal* 71 (282): 305–308.
Field, Alexander J. 2011. *A Great Leap Forward: 1930s Depression and Economic Growth*. New Haven, Conn.: Yale University Press.
Filippettia, Andrea, and Daniele Archibugia. 2010. "Innovation in Times of Crisis: National Systems of Innovation, Structure, and Demand." *Research Policy* 40 (2): 179–192.
Fink, Carsten, and Keith E. Maskus, eds. 2005. *Intellectual Property and Development*. Oxford: Oxford University Press.
Fleisher, Belton, Haizheng Li, and Min Qiang Zhao. 2010. "Human Capital, Economic Growth, and Regional Inequality in China." *Journal of Development Economics* 92 (2): 215–231.
Florida, Richard. 2002. *The Rise of the Creative Class: And How It's Transforming Work, Leisure, Community and Everyday Life*. New York: Perseus.
Foster, Lucia, John. C. Haltiwanger, and C. J. Krizan. 2001. "Aggregate Productivity Growth. Lessons from Microeconomic Evidence." In *New Developments in Productivity Analysis*, 303–372. Cambridge, Mass.: National Bureau of Economic Research.
Frankel, Jeffrey A., and David Romer. 1999. "Does Trade Cause Growth?" *American Economic Review* 89: 379–399.
Freeman, Christopher. 1995. "The National System of Innovation in Historical Perspective." *Cambridge Journal of Economics* 19: 5–24.
———. 1987. *Technology Policy and Economic Performance: Lessons from Japan*. London, New York: Pinter.

Friedland, Roger, and Robert R. Alford. 1991. "Bringing Society Back In: Symbols, Practices, and Institutional Contradictions." In *The New Institutionalism in Organizational Analysis*, ed. Walter W. Powell and Paul J. DiMaggio, 232–263. Chicago: University of Chicago Press.

Fryer, Roland, and Matthew O. Jackson. 2008. "A Categorical Model of Cognition and Biased Decision Making." *The B. E. Journal of Theoretical Economics* 8 (1), Article 6.

Fudenberg, Drew, and Jean Tirole. 1982. "Learning-by-Doing and Market Performance." CERAS D.P.8, Ecole Nationale des Ponts et Chaussees.

——. 1983. "Learning-by-Doing and Market Performance." *Bell Journal* 14 (2): 522–530.

Fudenberg, Drew, Rirchard Gilbert, Jean Tirole, and Joseph E. Stiglitz. 1983. "Preemption, Leapfrogging and Competition in Patent Races." *European Economic Review* 22 (June): 3–32.

Furman, J., and Joseph E. Stiglitz. 1999. "Economic Consequences of Income Inequality." In *Symposium Proceedings, 1998: Income Inequality: Issues and Policy Options*, 221–263. Symposium held by Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming.

——. 1998. Economic Crises: Evidence and Insights from East Asia. Brookings Papers on Economic Activity, no. 2, 1–114. Presented at Brookings Panel on Economic Activity, Washington, D.C., September 3, 1998.

Futia, Carl A. 1980. "Schumpeterian Competition." *The Quarterly Journal of Economics* 94 (4): 675–695.

Gallagher, Kevin. 2014. *Ruling Capital: Emerging Markets and the Reregulation of Cross-Border Finance*. Ithaca, New York: Cornell University Press.

Gallini, Nancy T. 2002. "The Economics of Patents: Lessons from Recent U.S. Patent Reform." *Journal of Economic Perspectives* 16: 131–154.

Gertner, Jon. 2012. *The Idea Factory: Bell Labs and the Great Age of American Innovation*. New York: Penguin.

Ghemawat, Pankaj, and A. Michael Spence. 1985. "Learning Curve Spillovers and Market Performance." *Quarterly Journal of Economics* 100: 839–852.

Gilbert, R. 2006. "Looking for Mr. Schumpeter: Where Are We in the Competition-Innovation Debate?" In Innovation Policy and the Economy, Vol. 6, eds. A. Jaffe, J. Lerner, S. Stern. Cambridge, Mass: MIT Press.

Gilbert, Richard J., and David M. Newbery. 1982. "Preemptive Patenting and the Persistence of Monopoly." *American Economic Review* 72 (3): 514–526.

Goettler, R. L. and Gordon, B. R. 2014. "Competition and Product Innovation in Dynamic Oligopoly." *Quant Mark Econ* 12: 1–42.

——. 2011. "Does AMD Spur Intel to Innovate More?" *Journal of Political Economy* 119(6) 1141–1200.

Goldin, Claudia, and Lawrence Katz. 2008. *The Race Between Education and Technology*. Cambridge, Mass.: Harvard University Press.

Goldstone, Lawrence. 2014. *Birdmen: The Wright Brothers, Glenn Curtiss, and the Battle to Control the Skies*. New York: Ballantine Books.

Goozner, Merrill. 2010. "Ruling on BRCA Gene Patents Could Have Limited Impact." *Journal of the National Cancer Institute* 102 (11): 754–757.

Gordon, R. 2012. "Is U.S. Economic Growth Over? Faltering Innovation Confronts the Six Headwinds." NBER Working Paper 18315.

Graham, Stuart, and Saurabh Vishnubhakat. 2013. "Of Smart Phone Wars and Software Patents." *Journal of Economic Perspectives* 27(1): 67–86.

Gramsci, Antonio. 1971. *Selections from the Prison Notebooks of Antonio Gramsci*. Ed. and trans. Quintin Hoare and Geoffrey Nowell Smith. New York: International.

Grandstrand, Ove. 2005. "Innovation and Intellectual Property Rights." In *The Oxford Handbook of Innovation*. Ed. I. Fagerberg, D. Mowery, and R. Nelson, 266–290. Oxford: Oxford University Press.

Greenwald, Bruce, and Judd Kahn. 2009. *Globalization: n. The Irrational Fear That Someone in China Will Take Your Job*. Hoboken, N. J.: John Wiley.

Greenwald, Bruce C., and Joseph E. Stiglitz. 2014a. "Industrial Policies, the Creation of a Learning Society, and Economic Development." In *The Industrial Policy Revolution I: The Role of Government Beyond Ideology*, ed. Joseph E. Stiglitz and Justin Yifu Lin, 43–71. New York: Palgrave Macmillan.

——. 2014b. "Learning and Industrial Policy: Implications for Africa." In *The Industrial Policy Revolution II: Africa in the 21st Century*, ed. Joseph E. Stiglitz, Justin Yifu Lin, and Ebrahim Patel, 25–29. New York: Palgrave Macmillan.

——. 2014c. *Creating a Learning Society: A New Paradigm for Development and Social Progress*. New York: Columbia University Press.

——. 2010a. "A Modest Proposal for International Monetary Reform." In *Time for a Visible Hand: Lessons from the 2008 World Financial Crisis*, ed. S. Griffith-Jones, J. A. Ocampo, and Joseph E. Stiglitz, 314–344. Initiative for Policy Dialogue Series. Oxford: Oxford University Press.

——. 2010b. "Towards a New Global Reserves System." *Journal of Globalization and Development* 1 (2), Article 10. A different version of the paper, with the same title, appears in *The Future Global Reserve System: An Asian Perspective*. Ed. J. D. Sachs, M. Kawai, J.-W. Lee, and W. T. Woo. Asian Development Bank, June.

——. 2006. "Helping Infant Economies Grow: Foundations of Trade Policies for Developing Countries." *American Economic Review: AEA Papers and Proceedings* 96 (2): 141–146.

——. 2003. *Towards a New Paradigm in Monetary Economics*. Cambridge, UK: Cambridge University Press.

——. 1993. "Financial Market Imperfections and Business Cycles." *Quarterly Journal of Economics* 108 (1): 77–114. Reprinted in *Selected Works of Joseph E. Stiglitz*, vol. 1, *Information and Economic Analysis*, 617–648. Oxford: Oxford University Press, 2009.

——. 1990. "Asymmetric Information and the New Theory of the Firm: Financial Constraints and Risk Behavior." *American Economic Review* 80 (2): 160–165. Also NBER Working Paper 3359.

——. 1988. "Pareto Inefficiency of Market Economies: Search and Efficiency Wage Models." *American Economic Review* 78 (2): 351–355.

——. 1986. "Externalities in Economies with Imperfect Information and Incomplete Markets." *Quarterly Journal of Economics* 1 (2): 229–264.

Greenwald, Bruce C., Alec Levinson, and Joseph E. Stiglitz. 1993. "Capital Market Imperfections and Regional Economic Development." In *Finance and Development: Issues and Experience*, ed. Alberto Giovannini, 65–93. Cambridge: Cambridge University Press.

Greenwald, B., M. Salinger, and J. E. Stiglitz. 1990. "Imperfect Capital Markets and Productivity Growth." Paper presented at NBER Conference in Vail, Colorado, April 1990, revised March 1991 and April 1992.

Greenwald, Bruce C., Joseph E. Stiglitz, and Andrew Weiss. 1984. "Informational Imperfections in the Capital Markets and Macroeconomic Fluctuations." *American Economic Review* 74 (2): 194–199.

Greif, Avner, and David Laitin. 2007. "A Theory of Endogenous Institutional Change." *American Political Science Review* 98 (4): 633–652.

Greiner, Alfred, Jens Rubart, and Willi Semmler. 2003. Economic Growth, Skill-Biased Technical Change and Wage Inequality: A Model and Estimations for the U.S. and Europe. New School Working Paper. http://www.newschool.edu/nssr/cem/papers/wp/labor/wineq.pdf.

Griliches, Zvi, and Dale W. Jorgenson. 1967. "The Explanation of Productivity Change." *Review of Economic Studies* 34 (3): 249–283.

——. 1966. "Sources of Measured Productivity Change: Capital Input." *American Economic Review* 56 (2): 50–61.

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

Grossman, S., and Joseph E. Stiglitz. 1980. "On the Impossibility of Informationally Efficient Markets." *American Economic Review* 70 (3): 393–408.

——. 1976. "Information and Competitive Price Systems." *American Economic Review* 66 (2): 246–253.

Gruber, Harald. 1998. "Learning by Doing and Spillovers: Further Evidence for the Semiconductor Industry." *Review of Industrial Organization* 13 (6): 697–711.

Guzman, Martin, Jose Antonio Ocampo, and Joseph E. Stiglitz. 2014. "Real Exchange Rate Policies for Economic Development." Initiative for Policy Dialogue, Columbia University (paper prepared for South African government).

Habakkuk, H. J. 1962. *American and British Technology in the Nineteenth Century*. London: Cambridge University Press.

Hagiu, Andrei, and David B. Yoffie. 2013. "The New Patent Intermediaries: Platforms, Defensive Aggregators, and Super-Aggregators." *Journal of Economic Perspectives* 27(1): 45–66.

Hahn, Frank. 1966. "Equilibrium Dynamics with Heterogeneous Capital Goods." *Quarterly Journal of Economics* 80: 633–646.

Hall, Bronwyn H. 1992. "Investment and Research and Development at the Firm Level: Does the Source of Financing Matter?" Working Paper, University of California, Berkeley.

——. 1991. "Corporate Restructuring and Investment Horizons." In *Capital Choices: Changing the Way America Invests in Industry*, ed. Michael Porter. Cambridge, Mass.: Harvard University Press.

——. 1990. "The Impact of Corporate Restructuring on Industrial Research and Development." *Brookings Papers on Economic Activity* 85–136.

Hall, Robert E., and Dale W. Jorgenson. 1967. "Tax Policy and Investment Behavior." *American Economic Review* 57 (June): 391–414.

Harberger, Arnold C. "Monopoly and Resource Allocation." *The American Economic Review*, 1954, *44* (Papers and Proceedings of the Sixty-sixth Annual Meeting of the American Economic Association): 7787.

——. 1971. "On Measuring the Social Opportunity Cost of Labour." *International Labor Review* 103 (6): 559–579.

Hart, Oliver. 1983. "The Market Mechanism as an Incentive Scheme." *Bell Journal of Economics,* 366–382.

Hausmann, Ricardo, and Bailey Klinger. 2007. "The Structure of the Product Space and the Evolution of Comparative Advantage." CID Working Paper No. 146. Center for International Development at Harvard University, April 2007. Available at http://www.hks.harvard.edu/var/ezp_site/storage/fckeditor/file/pdfs/centers-programs/centers/cid/publications/faculty/wp/146.pdf.

Hausmann, Ricardo, and Dani Rodrik. 2003. "Economic Development As Self-Discovery." *Journal of Development Economics* 72 (2): 603–633.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik. 2007. "What You Export Matters." *Journal of Economic Growth* 12 (1): 1–25.

Hayek, Friedrich A. 1945. "The Use of Knowledge in Society." *American Economic Review* 35 (4): 519–530.

Heilbroner, Robert L. 1980. *The Worldly Philosophers* 5th ed. New York: Simon and Schuster.

Helleiner, Gerald K. 1994. *Trade Policy and Industrialization in Turbulent Times*. New York: Routledge.

Heller, Michael. 2008. *The Gridlock Economy: How Too Much Ownership Wrecks Markets, Stops Innovation, and Costs Lives*. New York: Basic Books.

——. 1998. "The Tragedy of the Anticommons: A Concise Introduction and Lexicon." *Harvard Law Review* 76 (1): 6–25.

Heller, Michael A., and Rebecca S. Eisenberg. 1998. "Can Patents Deter Innovation? The Anti-Commons in Biomedical Research." *Science* 280: 698–701.

Hellman, Thomas, Kevin Murdock, and Joseph E. Stiglitz. 2002. "Franchise Value and the Dynamics of Financial Liberalization." In *Designing Financial Systems in Transition Economies: Strategies for Reform in Central and Eastern Europe*, ed. Anna Meyendorff and Anjan Thakor, 111–127. Cambridge, Mass.: MIT Press.

——. 2000. "Liberalization, Moral Hazard in Banking and Prudential Regulation: Are Capital Requirements Enough?" *American Economic Review* 90 (1): 147–165. Also published in *Industrial Organization and Regulation* 3 (17), August 2000.

——. 1998. "Financial Restraint and the Market Enhancing View." In *The Institutional Foundations of East Asian Economic Development*, ed. Y. Hayami and M. Aoki, 255–284. London: MacMillan.

——. 1997. "Financial Restraint: Toward a New Paradigm." In *The Role of Government in East Asian Economic Development*, ed. M. Aoki, H. Kim, and M. Okuna-Fujiwara, 163–207. Oxford: Clarendon Press.

——. 1996. "Deposit Mobilisation Through Financial Restraint." In *Financial Development and Economic Growth*, ed. N. Hermes and R. Lensink, 219–246. New York: Routledge.

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

Henry, Claude, and Joseph E. Stiglitz. 2010. "Intellectual Property, Dissemination of Innovation, and Sustainable Development." *Global Policy* 1 (1): 237–251.

Hertel, G., M. Krishnan, and S. Slaughter. 2003. "Motivation in Open Source Projects: An Internet-based Survey of Contributors to the Linux Kernel." *Research Policy* 32(7): 1159–1177.

Hertel, Guido, Stefanie Herrmann, and Sven Niedner. 2003. "Motivation in Open Source Projects: An Internet-based Survey of Contributors to the Linux Kernel." *Research Policy* 32 (7): 1159–1177.

Hicks, John R. 1935. "Annual Survey of Economic Theory: The Theory of Monopoly." *Econometrica* 3 (1): 1–20.

——. 1932. *Theory of Wages*. London: Macmillan; New York: St. Martin's.

Hidalgo, César A., and Ricardo Hausmann. 2009. "The Building Blocks of Economic Complexity." *Proceedings of the National Academy of Sciences of the United States of America*, 106 (26), 10570–5.

Hidalgo, César A., Bailey Klinger, Albert-László Barabási, and Ricardo Hausmann. 2007. "The Product Space Conditions the Development of Nations." *Science* 317 (5837, July): 482–487.

Hirsch, Werner. 1952. "Manufacturing Progress Functions." *Review of Economics and Statistics* 34: 143–155.

Hirschman, Albert O. 1982. "The Rise and Decline of Development Economics." In *The Theory and Experience of Economic Development*, ed. M. Gersovitz and W. A. Lewis, 372–390. London: Allen and Unwin.
——. 1958. *The Strategy of Economic Development*. New Haven, Conn.: Yale University Press.
Hoff, Karla. 1997. "Bayesian Learning in an Infant Industry Model." *Journal of International Economics* 43 (3–4): 409–436.
Hoff, Karla, and Priyanka Pandey. 2011. Names Can Hurt You (by Cueing Your Response to Incentives): Experimental Evidence on Identity and Development. Manuscript, World Bank.
——. 2006. "Discrimination, Social Identity, and Durable Inequalities." *American Economic Review, Papers and Proceedings* 96: 206–211.
——. 2005. "Opportunity Is Not Everything: How Belief Systems and Mistrust Shape Responses to Economic Incentives." *Economics of Transition* 13 (2), special issue, *Institutions and Economic Performance*: 445–472.
Hoff, Karla, and Arijit Sen. 2006. "The Kin System as a Poverty Trap?" In *Poverty Traps*, ed. Samuel Bowles, Steven Durlauf, and Karla Hoff, 95–115. Princeton, N.J.: Princeton University Press.
Hoff, Karla, and Joseph E. Stiglitz. 2011. "The Role of Cognitive Frames in Societal Rigidity and Change." World Bank. http://www.econ.yale.edu/seminars/develop/tdw11/hoff-110404a.pdf.
——. 2010. "Equilibrium Fictions: A Cognitive Approach to Societal Rigidity." *American Economic Review* 100 (2, May): 141–146. Extended version available as Policy Research Working Paper 5219, World Bank Development Research Group, February 2010. http://www-wds.worldbank.org/external/default/WDSContentServer/IW3P/IB/2010/02/26/000158349_20100226083837/Rendered/PDF/WPS5219.pdf.
——. 2007. "Exiting a Lawless State." *Economic Journal* 118 (531): 1474–1497.
——. 2004a. "After the Big Bang? Obstacles to the Emergence of the Rule of Law in Post-Communist Societies." *American Economic Review* 94 (3): 753–763.
——. 2004b. "The Transition Process in Post-Communist Societies: Towards a Political Economy of Property Rights." In *Toward Pro-Poor Policies: Aid, Institutions and Globalization*, ed. B. Tungodden, N. Stern, and I. Kolstad, 231–245. New York: World Bank/Oxford University Press.
——. 2001. "Modern Economic Theory and Development." In *Frontiers of Development Economics: The Future in Perspective*, ed. G. Meier and Joseph Stiglitz, 389–459. Oxford: Oxford University Press.
Hollander, Samuel. 1965. *The Sources of Increased Efficiency: A Study of Du Pont Rayon Plants*. Cambridge, Mass.: MIT Press.
Holt, Jeff. 2009. "A Summary of the Primary Causes of the Housing Bubble and the Resulting Credit Crisis: A Non-Technical Paper." *Journal of Business Inquiry* 8 (1): 120–129.

Honohan, Patrick, and Joseph E. Stiglitz. 2001. "Robust Financial Restraint." In *Financial Liberalization: How Far, How Fast?* ed. G. Caprio, P. Honohan, and Joseph Stiglitz, 31–63. Cambridge, UK: Cambridge University Press.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics* 124 (4): 1403–1448.

Huang, Kenneth G. and Fiona E. Murray. 2008. "Does Patent Strategy Shape the Long Run Supply of Public Knowledge? Evidence from Human Genetics." *Academy of Management Journal* 52(6): 1193–1221.

Humphreys, Macartan, Jeffrey D. Sachs, and Joseph E. Stiglitz, eds. 2007. *Escaping the Resource Curse*. New York: Columbia University Press.

Im, F. G. and D. Rosenblatt. 2013. Middle-Income Traps: A Conceptual and Empirical Survey. Policy Research Working Paper 6594, Washington, D.C.: World Bank.

IMF. 2012a. "Sovereigns, Banks, and Emerging Markets: Detailed Analysis and Policies." *Global Financial Stability Report*, chapter 2. http://www.imf.org/external/pubs/ft/gfsr/2012/01/pdf/c2.pdf.

IMF. 2012b. "The Liberalization and Management of Capital Flows: An Institutional View." Online publication of the International Monetary Fund. http://www.imf.org/external/np/pp/eng/2012/111412.pdf.

Inada, Ken-Ichi, 1963. "On a Two-Sector Model of Economic Growth: Comments and a Generalization," *The Review of Economic Studies*, 30(2): 119–127.

Irwin, Douglas A., and Randall S. Kroszner. 1999. "Interests, Institutions, and Ideology in Securing Policy Change: The Republican Conversion to Trade Liberalization after Smoot-Hawley." *Journal of Law and Economics* 42 (2, October): 643–674.

Ismail, Faizel. 2007. "Mainstreaming Development in the WTO: Developing Countries in the Doha Round." Report of CUTS International, available online at http://library.fes.de/pdf-files/bueros/genf/04888.pdf.

Jaffe, Adam, and Josh Lerner. 2004. *Innovation and Its Discontents*. Princeton, N. J.: Princeton University Press.

Jaffe, Adam, Josh Lerner, and Scott Stern, eds. Annual. *Innovation Policy and the Economy*. Cambridge, Mass.: MIT Press for National Bureau of Economic Research.

Jaffee, Dwight, and Joseph E. Stiglitz. 1990. "Credit Rationing." In *Handbook of Monetary Economics*, ed. B. Friedman and F. Hahn, 837–888. Amsterdam: Elsevier Science Publishers.

Janeway, William. 2012. *Doing Capitalism in the Innovation Economy*. Cambridge, Mass., Cambridge University Press

Jarmin, Ron S. 1994. "Learning by Doing and Competition in the Early Rayon Industry." *RAND Journal of Economics* 25 (3): 441–454.

Jayadev, Arjun, and Joseph E. Stiglitz. 2010. “Medicine for Tomorrow: Some Alternative Proposals to Promote Socially Beneficial Research and Development in Pharmaceuticals.” *Journal of Generic Medicines* 7 (3): 217–226.

——. 2009. “Two Ideas to Increase Innovation and Reduce Pharmaceutical Costs and Prices.” *Health Affairs* 28 (1): w165–w168.

Johnson, Simon, and James Kwak. 2010. *13 Bankers: The Wall Street Takeover and the Next Financial Meltdown*. New York: Vintage.

Jourdan, P. 2014. “Towards a Resource-Based African Industrialisation Strategy.” *The Industrial Policy Revolution II: Africa in the 21st Century*, ed. Joseph E. Stiglitz, Justin Yifu Lin, and Ebrahim Patel, 364–386. New York: Palgrave Macmillan.

Jovanovic, Boyan, and Saul Lach. 1989. “Entry, Exit, and Diffusion with Learning by Doing.” *American Economic Review* 79 (4): 690–699.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Kaldor, Nicholas. 1961. “Capital Accumulation and Economic Growth.” In *The Theory of Capital*, ed. F. A. Lutz and D. C. Hague, International Economic Association Conference. London: Macmillan; New York: St. Martin’s.

——. 1957. “A Model of Economic Growth.” *Economic Journal* 67:591–624.

——. 1934. “The Equilibrium of the Firm.” *Economic Journal* 44 (173): 60–76.

Kaldor, Nicholas, and James A. Mirrlees. 1962. “A New Model of Economic Growth.” *Review of Economic Studies* 29 (June): 174–192.

Kamien, Mort, and Nancy Schwartz. 1975. “Market Structure and Innovation: A Survey.” *Journal of Economic Literature* 13(1): 1–37.

——. 1972. “Timing of Innovation and Rivalry,” *Econometrica*, 40 (1): 43–60.

Kanbur, R. 1979. “Impatience, Information, and Risk Taking in a General Equilibrium Model of Occupational Choice.” *Review of Economic Studies* 46 (4): 707–718.

Kaplan, S., and J. Lerner. 2010. “It Ain't Broke: The Past, Present, and Future of Venture Capital.” *Journal of Applied Corporate Finance* 22(2): 36–47.

Karl, Terry Lynn. 1997. *The Paradox of Plenty: Oil Booms and Petro-States*. Berkeley, California: University of California Press.

Keller, Wolfgang. 2004. “International Technology Diffusion.” *Journal of Economic Literature* 42: 752–782.

——. 2002. “Technology Diffusion and the World Distribution of Income: The Role of Geography, Language, and Trade.” University of Texas, unpublished.

Kennedy, Charles. 1964. “Induced Bias in Innovation and the Theory of Distribution.” *Economic Journal* 74 (September): 541–547.

Keynes, John Maynard. 1930. “Economic Possibilities for Our Grandchildren.” In *Essays in Persuasion*, 358–373 New York: W. W. Norton, 1963.

Khan, Mushtaq H. 2014. "Technology Policies and Learning with Imperfect Governance." In *The Industrial Policy Revolution II: Africa in the 21st Century*, ed. Joseph E. Stiglitz, Justin Yifu Lin, and Ebrahim Patel, 243–280. New York: Palgrave Macmillan.

——. 2012. "Governance and Growth: History, Ideology and Methods of Proof." In *Good Growth and Governance in Africa: Rethinking Development Strategies*, ed. Akbar Noman, Kwesi Botchwey, Howard Stein, and Joseph E. Stiglitz. 51–79. New York: Oxford University Press.

Kihlstrom, Richard E., and Jean Jacques Laffont. 1979. "A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion." *Journal of Political Economy* 87: 719–848.

Kindleberger, Charles P., and Robert Aliber. 2005. *Manias, Panics, and Crashes: A History of Financial Crises*, 5th ed. New York: Wiley Investment Classics.

Klenow, Peter J., and Andrés Rodríguez-Clare. 1997. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" In *NBER Macroeconomics Annual 1997*, vol. 12, ed. Ben S. Bernanke and Julio J. Rotemberg, 73–103. Cambridge, Mass.: MIT Press.

Knight, Frank. 1921. *Risk, Uncertainty, and Profit*. Boston: Houghton Mifflin.

Korinek, Anton, and Luis Servén. 2010. "Undervaluation Through Foreign Reserve Accumulation: Static Losses, Dynamic Gains." Policy Research Working Paper 5250, Washington, D.C.: World Bank.

Kose, M. Ayhan, Eswar Prasad, Kenneth Rogoff, and Shang-Jin Wei. 2006. "Financial Globalization: A Reappraisal." IMF Working Papers, WP/06/189.

Kremer, Michael. 1998. "Patent Buy-outs: A Mechanism for Encouraging Innovation." *Quarterly Journal of Economics* (November): 1137–1167.

Kremer, Michael, and H. Williams. 2010. "Incentivizing Innovation: Adding to the Tool Kit." In *Innovation Policy and the Economy*, ed. J. Lerner and S. M. Stern, 1–17. Cambridge, Mass.: MIT Press for the National Bureau of Economic Research.

Krueger, Anne O., and Baran Tuncer. 1982. "An Empirical Test of the Infant Industry Argument." *American Economic Review* 72: 1142–1152.

Krugman, Paul. 1981. "Trade, Accumulation and Uneven Development." *Journal of Development Economics* 8 (2): 149–161.

——. 1979. "Increasing Returns, Monopolistic Competition, and International Trade." *Journal of International Economics* 9 (4): 469–479.

Leahy, Dermot, and J. Peter Neary. 1999. "Learning by Doing, Precommitment and Infant-Industry Promotion." *Review of Economic Studies* 66 (2): 447–474.

Lederman, Daniel, Ana Maria Menéndez, Guillermo Perry, and Joseph Stiglitz. 2003. "Mexican Investment After the Tequila Crisis: Basic Economics, 'Confidence' Effects or Market Imperfections?" *Journal of International Money and Finance* 22: 131–151.

Leibenstein, Harvey. 1966. “Allocative Efficiency vs. X-Efficiency.” *American Economic Review* 56 (3): 392–415.

Lemley, Mark A., and Carl Shapiro. 2007. “Patent Holdup and Royalty Stacking.” *Texas Law Review* 85(7): 2163–2173.

Lerner, Josh, and Jean Tirole. 2002. “Some Simple Economics of Open Source.” *Journal of Industrial Economics* 50 (2): 197–234.

Levhari, David. 1967. “Further Implications of Learning by Doing.” *Review of Economic Studies* 33: 31–38.

——. 1966. “Extensions of Arrow’s Learning by Doing.” *Review of Economic Studies* 33 (2): 31–38.

Levin, Richard C. 1978. “Technical Change, Barriers to Entry, and Market Structure.” *Economica* 45 (180): 347–362.

Levin, Richard C., Alvin K. Klevorick, Richard. R. Nelson, and Sidney Winter. 1987. “Appropriating the Returns from Industrial R & D.” *Brookings Papers on Economic Activity* (3): 783–831.

Lewis, Tracy, and Jerome H. Reichman. 2005. “Using Liability Rules to Stimulate Local Innovation in Developing Countries: Application to Traditional Knowledge.” In *International Public Goods and Transfer of Technology under a Globalized Intellectual Property Regime*, ed. K. Maskus and J. Reichman, 337–366. Cambridge, Mass.: Cambridge University Press.

Lieberman, Marvin B. 1987. “The Learning Curve, Diffusion, and Competitive Strategy.” *Strategic Management Journal* 8 (5): 441–452.

——. 1984. “The Learning Curve and Pricing in the Chemical Processing Industries.” *RAND Journal of Economics* 15: 213–228.

Lin, Justin Y. “From Flying Geese to Leading Dragons: New Opportunities and Strategies for Structural Transformation in Developing Countries,” in *The Industrial Policy Revolution II: Africa in the 21st Century*, ed. Joseph E. Stiglitz, Justin Yifu Lin, and Ebrahim Patel, 50–70. New York: Palgrave Macmillan.

——. 2012. *New Structural Economics: A Framework for Rethinking Development and Policy*. Washington, D.C.: World Bank.

——. 2010. “New Structural Economics: A Framework for Rethinking Development.” Policy Research Working Paper 5197.

Lin, Justin Y., and Célestin Monga. 2014. “Comparative Advantage: The Silver Bullet of Industrial Policy.” In The Indistrial Policy Revolution I: The Role of Government Beyond Ideology, ed. Joseph E. Stiglitz and Justin Yifu Lin, 19–38. New York: Palgrave Macmillan.

Lin, Justin Y., Fang Cai, and Zhou Li. 2003. *The China Miracle: Development Strategy and Economic Reform*. Sha Tin, N.T., Hong Kong: Chinese University Press for the Hong Kong Centre for Economic Research and the International Center for Economic Growth. Lin paper is in the IEA Washington volume.

Lipsey, R. G., and Kelvin Lancaster. 1956–1957. "The General Theory of Second Best." *Review of Economic Studies* 24 (1): 11–32.
List, Friedrich. 1841. *Das nationale System der politischen Oekonomie*. Stuttgart: J. G. Cotta.
Loury, Glenn C. 2002. *The Anatomy of Racial Inequality*. Cambridge, Mass.: Harvard University Press.
——. 1979. "Market Structure and Innovation." *Quarterly Journal of Economics* 93: 395–410.
Love, James. 2004. "Compensation Guidelines for Non-Voluntary Use of a Patent on Medical Technologies." Working Paper.
Love, James, and T. Hubbard. 2007. "The Big Idea: Prizes to Stimulate R & D for New Medicines." *Chicago-Kent Law Review* 1519–1554.
Lucas, Robert E. Jr. 1993. "Making a Miracle." *Econometrica* 61 (2): 251–272.
——. 1990. "Why Doesn't Capital Flow from Rich to Poor Countries?" *American Economic Review* 80 (2): 92–96.
——. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22: 3–42.
Lundvall, Bengt-Åke, ed. 2010. *National Systems of Innovation: Toward a Theory of Innovation and Interactive Learning*. London: Anthem.
Maddison, A. 2001. *The World Economy: A Millennial Perspective*. Paris: Development Centre of the Organisation for Economic Co-operation and Development.
Malerba, Franco. 1992. "Learning by Firms and Incremental Technical Change." *Economic Journal* 102 (413): 845–859.
Malhotra, Kamal. 2003. *Making Global Trade Work for People*. London: Earthscan.
Mankiw, G., D. Romer, and D. Weil. 1992. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* 107(2): 407–437.
Mansfield, Edwin. 1967. *Econometric Studies of Industrial Research and Technological Innovation*. New York: Norton.
Martin, Stephan. 2000. "The Theory of Contestable Markets." Working Paper, Purdue University. http://www.krannert.purdue.edu/faculty/smartin/aie2/contestbk.pdf.
Matsuyama, Kiminori. 1992. "Agricultural Productivity, Comparative Advantage, and Economic Growth." *Journal of Economic Theory* 58(2): 317–334.
Mazzucato, Mariana. 2013. *The Entrepreneurial State: Debunking Public vs. Private Sector Myths*. London: Anthem.
McCall, Morgan W. Jr. 2004. "Leadership Development Through Experience." *Academy of Management Executive* 18 (3): 127–130.
Meade, James E. 1955. *Trade and Welfare*. London: Oxford University Press.
Merges, R., and R. Nelson. 1994. "On Limiting or Encouraging Rivalry in Technical Progress: The Effects of Patent Scope Decisions." *Journal of Economic Behavior and Organization* 25: 1–24.

Merton, Robert. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago Press.

Mian, Atif. 2006. "Distance Constraints: The Limits of Foreign Lending in Poor Economies." *Journal of Finance* 61 (3): 1465–1505.

Mokyr, Joel. 2009. *The Enlightened Economy: An Economic History of Britain, 1700–1850*. New Haven, Conn.: Yale University Press.

Mookherjee, Dilip, and Debraj Ray. 2003. "Persistent Inequality." *Review of Economic Studies* 70: 369–393.

Moretti, Enrico. 2011. "Local Labor Markets." In *Handbook of Labor Economics*, ed. O. Ashenfelter and D. E. Card, 1237–1313. Amsterdam: Elsevier.

Moser, Petra. 2013. "Patents and Innovation: Evidence from Economic History." *Journal of Economic Perspectives* 27 (1): 23–44.

Mowery, David C., Richard R. Nelson, Bhaven N. Sampat, and Arvids A. Ziedonis. 2001. "The Growth of Patenting and Licensing by U.S. Universities: An Assessment of the Effects of the Bayh–Dole Act of 1980." *Research Policy* 30 (1): 99–119.

Murdock, K., and Joseph E. Stiglitz. 1993. "The Effect of Financial Repression in an Economy with Positive Real Interest Rates: Theory and Evidence." Working Paper.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny. 1989. "Industrialization and the Big Push." *Journal of Political Economy* 97 (5): 1003–1026.

Myrdal, Gunnar. 1968. *Asian Drama: An Inquiry into the Poverty of Nations*. New York: Twentieth Century Fund.

Nakamura, Maso, Sadao Sakakibara, and Roger Schroeder. 1998. "The Adoption of Just-in-Time Manufacturing Methods at US- and Japanese-Owned Plants: Some Empirical Evidence." *IEEE Transactions on Engineering Management* 45 (3): 230–240.

Nalebuff, Barry J., and Joseph E. Stiglitz. 1983a. "Information, Competition and Markets." *American Economic Review* 73 (2, May): 278–284. Reprinted in *Selected Works of Joseph E. Stiglitz*, vol. 1, *Information and Economic Analysis*, 400–408. Oxford: Oxford University Press, 2009.

——. 1983b. "Prizes and Incentives: Toward a General Theory of Compensation and Competition." *Bell Journal* 14 (1, Spring): 21–43. Reprinted in *Selected Works of Joseph E. Stiglitz*, vol. 2, *Information and Economic Analysis: Applications to Capital, Labor, and Product Markets*, 407–431. Oxford: Oxford University Press, 2013.

National Commission on the Causes of the Financial and Economic Crisis in the United States. 2011. *The Financial Crisis Inquiry Report of the Bipartisan National Commission on the Causes of the Financial and Economic Crisis in the United States*.

Nelson, Richard R. 2004. "The Challenge of Building an Effective Innovation System for Catch-up." *Oxford Development Studies* 32 (3): 365–374.

Nelson, Richard R., and Edmund S. Phelps. 1965. "Investment in Humans, Technological Diffusion and Economic Growth." Discussion Paper 189, Cowles Foundation, Yale University, New Haven, Conn ecticut.
Nelson, R. R., and Sidney G. Winter. 1993. "In Search of Useful Theory of Innovation." *Research Policy* 22 (2): 108.
——. 1982. *An Evolutionary Theory of Economic Change*. Cambridge: Belknap Press of Harvard University Press.
——. 1977. "Forces Generating and Limiting Concentration under Schumpeterian Competition." *Bell Journal* 9 (2): 524–548.
Newbery, David, and Joseph E. Stiglitz. 1982. "The Choice of Techniques and the Optimality of Market Equilibrium with Rational Expectations." *Journal of Political Economy* 90 (2): 223–246.
Noman, Akbar, and Joseph E. Stiglitz. 2012a. "African Development Prospects and Possibilities." In *The Oxford Companion to the Economics of Africa*, ed. E. Aryeetey et al., 33–40. Oxford: Oxford University Press.
——. 2012b. "Introduction and Overview." *Good Growth and Governance for Africa: Rethinking Development Strategies*, ed. A. Noman, K. Botchwey, H. Stein, and Joseph E. Stiglitz. New York: Oxford University Press.
——. 2012c. "Strategies for African Development." In *Good Growth and Global Governance in Africa*, ed. A. Noman, K. Botchwey, H. Stein, and Joseph E. Stiglitz, 3–47. New York: Oxford University Press.
Noman, Akbar, K. Botchwey, H. Stein, and Joseph E. Stiglitz, eds. 2012. *Good Growth and Governance for Africa: Rethinking Development Strategies*. New York: Oxford University Press.
Nordhaus, William D. 1969a. "An Economic Theory of Technological Change." *American Economic Association Papers and Proceedings* 59 (May): 18–28.
——. 1969b. *Invention, Growth and Welfare: A Theoretical Treatment of Technological Change*. Cambridge, Mass.: MIT Press.
North, Douglass C. 2005. *Understanding the Process of Economic Change*. Princeton, N.J.: Princeton University Press.
——. 1991. "Institutions." *Journal of Economic Perspectives* 5 (1): 97–112.
——. 1971. "Institutional Change and Economic Growth." *Journal of Economic History* 31 (1): 118–125.
North, Douglass, and Robert Thomas. 1973. *The Rise of the Western World*. Cambridge: Cambridge University Press.
Nunn, Nathan, and Daniel Trefler. 2010. "The Structure of Tariffs and Long-Term Growth." *American Economic Journal: Macroeconomics* 2(4): 158–194.
Obstfeld, Maurice, and Kenneth Rogoff. 2000. "The Six Major Puzzles in International Macroeconomics: Is There a Common Cause?" In *NBER Macroeconomics Annual 2000*, ed. B. S. Bernanke and K. Rogoff, 339–390. Cambridge, Mass.: MIT Press.

Ocampo, José Antonio, and Bilge Erten. 2014. "Capital Account Regulations: Role, Effectiveness, and the IMF's Institutional View." In *Financial Stability and Growth: Essays on Financial Regulation and New Developmentalism*, ed. Luiz Carlos Bresser-Pereira, Jan Kregel, and Leonardo Burlamaqui. Milton Park: Routledge.

Odagiri Hiroyuki, Akira Goto, Atsushi Sunami, and Richard R. Nelson, eds. 2010. *Intellectual Property Rights, Development, and Catch Up: An International Comparative Study*. Oxford: Oxford University Press.

OECD. 2011. *OECD Factbook*. Data available online via Google Public Data explorer, http://www.google.com/publicdata/directory.

——. 2009. "Policy Responses to the Economic Crisis: Investing in Innovation for Long-Term Growth." June. Report published online at http://www.oecd.org/science/inno/42983414.pdf.

Ohlin, B. 1933. *Interregional and International Trade*. Cambridge, MA: Harvard University Press.

Ostry, J. D., A. Berg, and C. G. Tsangarides. 2014. Redistribution, Inequality and Growth. International Monetary Fund Discussion Paper.

Patel, Pari, and K. Pavitt. 1994. "The Nature and Economic Importance of National Innovation Systems." *STI Review* 14.

Pecchi, Lorenzo, and Gustavo Piga, eds. 2008. *Revisiting Keynes: Economic Possibilities for Our Grandchildren*. Cambridge, Mass.: MIT Press.

Perleman, Michael. 2002. *Steal This Idea: Intellectual Property and the Corporate Confiscation of Creativity*. New York: Palgrave.

Phelps, Edmund S. 2013. *Mass Flourishing: How Grassroots Innovation Created Jobs, Challenge, and Change*. Princeton, N.J.: Princeton University Press.

Phillips, P. C. B., and Perron, P. 1988. "Testing for a Unit Root in Time Series Regression." *Biometrika* 75 (2): 335–346.

Pietronero, Luciano, Matthieu Cristelli, and Andrea Tacchella. 2013. New Metrics for Economic Complexity: Measuring the Intangible Growth Potential of Countries. Paper presented at the Plenary Conference of the Institute for New Economic Thinking, Hong Kong, April 4–7.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Cambridge, Mass.: Harvard University Press.

——.1995. "Social Mobility and Redistributive Politics." *Quarterly Journal of Economics* 110: 551–583.

Piketty, T., Emmanuel Saez, and Stefanie Stantcheva. 2011. "Taxing the 1%: Why the Top Tax Rate Could Be Over 80%." *Vox* December 8.

Pollack, Andrew. 2011. "Ruling Upholds Gene Patent in Cancer Test." *New York Times*, July 29. http://www.nytimes.com/2011/07/30/business/gene-patent-in-cancer-test-upheld-by-appeals-panel.html.

Porter, Michael E. 1990. *The Competitive Advantage of Nations*. New York: Free Press. (Republished with a new introduction, 1998.)

Prasad, Eswar, Kenneth Rogoff, Shang-Jin Wei, and M. Ayhan Kose. 2003. The Effects of Financial Globalization on Developing Countries: Some Empirical Evidence. IMF Occasional Paper no. 220. Washington, D.C.: International Monetary Fund.

Rabin, Matthew, and Joel L. Schrag. 1999. "First Impressions Matter: A Model of Confirmatory Bias." *Quarterly Journal of Economics* (February): 37–82.

Radner, Roy, and Joseph E. Stiglitz. 1984. "A Nonconcavity in the Value of Information." In *Bayesian Models in Economic Theory*, ed. Marcel Boyer and Richard Kihlstrom, 33–52. Amsterdam: Elsevier Science Publications. Reprinted in *Selected Works of Joseph E. Stiglitz*, vol. 1, *Information and Economic Analysis*, 537–554. Oxford: Oxford University Press, 2009.

Rajan, Raghuram, and Luigi Zingales. 1998. "Financial Dependence and Growth." *American Economic Review* 88: 559–586.

Rashid, Hamid. 2012. "Foreign Banks, Competition for Deposits and Terms and Availability of Credit in Developing Countries." Working Paper.

——. 2011. "Credit to Private Sector, Interest Spread and Volatility in Credit-Flows: Do Bank Ownership and Deposits Matter?" DESA Working Paper No. 105, ST/ESA/2011/DWP/105. http://www.un.org/esa/desa/papers/2011/wp105_2011.pdf.

Reinganum, Jennifer F. 1982a. "A Dynamic Game of R & D: Patent Protection and Competitive Behavior." *Econometrica* 50 (3): 671–688.

——. 1982b. "Patent Races with a Sequence of Innovations." Working Paper, Cal Tech.

——. 1981. "Dynamic Games of Innovation." *Journal of Economic Theory* 25 (1): 21–41.

Rodriguez, Francisco, and Dani Rodrik. 2001. "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence." In *NBER Macroeconomics Annual 2000*, vol. 15, ed. Ben S. Bernanke and Kenneth Rogoff, 261–338. Cambridge, Mass.: MIT Press.

Rodrik, Dani. 2013. "Unconditional Convergence in Manufacturing." *Quarterly Journal of Economics* 128 (1): 165–204.

——. 2010. "Diagnostics Before Prescription." *Journal of Economic Perspectives* 24 (3, Summer): 33–44.

——. 2001. The Global Governance of Trade: As If Development Really Mattered. Background Paper for the Trade and Sustainable Human Development Project, United Nations Development Programme, New York.

Rodrik, Dani, and Arvind Subramanian. 2005. "From Hindu Growth to Productivity Surge: The Mystery of the Indian Growth Transition." *IMF Staff Papers* 52, no. 2.P.

Romer, Paul. 1994. "The Origins of Endogenous Growth." *Journal of Economic Perspectives* 8 (1): 3–22.

——. 1990. "Endogenous Technological Change." *Journal of Political Economy* 98: S71–S102.
——. 1986. "Increasing Returns and Long-Run Growth." *Journal of Political Economy* 94 (5, October): 1002–1037.
Rothschild, Michael, and Joseph E. Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics* 90 (4, November): 629–649.
Ruttan, Vernon W., and Yujiro Hayami. 1984. "Toward a Theory of Induced Institutional Innovation." *Journal of Development Studies* 20 (4): 203–223.
Sachs, Jeffrey D., and Andrew Warner. 1995. "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity*, Economic Studies Program, The Brookings Institution, vol. 26 (1, 25th A): 1–118.
Sah, Raj Kumar, and Joseph E. Stiglitz. 1991. "The Quality of Managers in Centralized Versus Decentralized Organizations." *Quarterly Journal of Economics* 106 (1, February): 289–325.
——. 1989a. "Sources of Technological Divergence Between Developed and Less Developed Countries." In *Debt, Stabilizations and Development: Essays in Memory of Carlos Diaz-Alejandro*, ed. G. Calvo, 423–446. Oxford: Basil Blackwell.
——. 1989b. "Technological Learning, Social Learning and Technological Change." In *The Balance Between Industry and Agriculture in Economic Development*, ed. S. Chakravarty, 285–298. London: Macmillan.
——. 1988a. "Committees, Hierarchies and Polyarchies." *Economic Journal* 98 (391, June): 451–470.
——. 1988b. "Qualitative Properties of Profit-Maximizing *K*-out-of-*N* Systems Subject to Two Kinds of Failure." *IEEE Transactions on Reliability* 37 (5, December): 515–520.
——. 1987a. "The Invariance of Market Innovation to the Number of Firms." *RAND Journal of Economics* 18 (1, Spring): 98–108.
——. 1987b. "Price Scissors and the Structure of the Economy." *Quarterly Journal of Economics* 102: 109–134.
——. 1986. "The Architecture of Economic Systems: Hierarchies and Polyarchies." *American Economic Review* 76 (4, September): 716–727.
——. 1985. "Human Fallibility and Economic Organization." *American Economic Review* 75 (2, May): 292–296. Reprinted in *Selected Works of Joseph E. Stiglitz*, vol. 1, *Information and Economic Analysis*, 409–416. Oxford: Oxford University Press, 2009.
Salter, W. E. G. 1966. *Productivity and Technical Change*, 2nd ed. Cambridge: Cambridge University Press.
Samuelson, Paul A. 1965. "A Theory of Induced Innovation along Kennedy-Weizsacker Lines." *Review of Economics and Statistics* 33: 133–146.

——. 1954. "The Pure Theory of Public Expenditure." *Review of Economics and Statistics* 36: 387–389.

——. 1948. "International Trade and the Equalisation of Factor Prices." *Economic Journal* 58 (230, June): 163–184.

——. 1938. "Welfare Economics and International Trade." *American Economic Review* 28: 261–266.

Sappington, David, and Joseph E. Stiglitz. 1987. "Information and Regulation." In *Public Regulation*, ed. E. Bailey, 3–43. London: MIT Press.

Scherer, Frederic M. 1967. "Market Structure and the Employment of Scientists and Engineers." *American Economic Review,* 524–531.

——. 1970. *Industrial Market Structure and Economic Performance*, Chicago: Rand McNally.

Schmookler, Jacob. 1962. "Economic Sources of Innovative Activity." *Journal of Economic History* 22: 1–20.

Schumpeter, Joseph A. 1951 [1934]. "Depressions: Can We Learn from Past Experience?" In *Essays of J. A. Schumpeter*, ed. R. V. Clemence, 108–117. Cambridge, Mass.: Addison-Wesley.

——. 1943. *Capitalism, Socialism and Democracy*. New York: Harper.

——. 1912. *Theorie der Wirtschaflichen Entwicklung*. Leipzig: Duncker and Humbolt. Translated in 1934 as *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*.

Schwoon, Malte. 2008. "Learning by Doing, Learning Spillovers and the Diffusion of Fuel Cell Vehicles." *Simulation Modelling Practice and Theory* 16 (9): 1463–1476.

Scitovsky, Tibor. 1986. *Human Desire and Economic Satisfaction: Essays on the Frontiers of Economics*. New York: New York University Press.

Scotchmer, Suzanne. 2004. *Innovation and Incentives*. Cambridge, Mass.: MIT Press.

Serra, Narcís, and Joseph E. Stiglitz, eds. 2008. *The Washington Consensus Reconsidered: Towards a New Global Governance*. New York: Oxford University Press.

Shapiro, Carl. 2010. "Injunctions, Hold-Up, and Patent Royalties." *American Law and Economics Review* 12 (2): 280–313.

——. 2007. "Patent Reform: Aligning Reward and Contribution." Working Paper 13141, National Bureau of Economic Research, Cambridge, Mass. http://papers.nber.org/papers/w13141.

——. 2001. "Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard Setting." In *Innovation Policy and the Economy*, ed. Adam Jaffe, Josh Lerner, and Scott Stern, 119–150. Cambridge, Mass.: MIT Press.

Shapiro, Carl, and Joseph E. Stiglitz. 1984. "Equilibrium Unemployment as a Worker Discipline Device." *American Economic Review* 74 (3): 433–444.

Shell, Karl. 1966. "Toward a Theory of Inventive Activity and Capital Accumulation." *American Economic Association Papers and Proceedings* 56: 62–68.
——, ed. 1967. *Essays on the Theory of Optimal Economic Growth*. Cambridge, Mass.: MIT Press.
Shell, Karl, and Joseph E. Stiglitz. 1967. "Allocation of Investment in a Dynamic Economy." *Quarterly Journal of Economics* 81 (November): 592–609.
Sheshinski, E. 1967a. "Optimal Accumulation with Learning by Doing." In *Essays on the Theory of Optimal Economic Growth*, ed. Karl Shell, 31–52. Cambridge, Mass.: MIT Press.
——. 1967b. "Tests of the Learning-by-Doing Hypothesis." *Review of Economics and Statistics* 49: 568–578.
Shierholz, Heidi, Lawrence Mishel, and John Schmitt. 2013. "Don't Blame the Robots: Assessing the Job Polarization Explanation of Growing Wage Inequality." *Economic Policy Institute* November 19.
Simon, Herbert A. 1991. "Organizations and Markets." *Journal of Economic Perspectives* 5 (2): 25–44.
Skeath, Susan. 1997. *Learning from "Learning by Doing": Lessons for Economic Growth*. Palo Alto, Calif.: Stanford University Press.
——. 1995. "A Role for Trade Policy? Markets with Informational Barriers to Entry." *International Trade Journal* 9 (2): 247–271.
——. 1993. "Strategic Product Choice and Equilibrium Traps for Less Developed Countries." *Journal of International Trade and Economic Development* 2 (1): 1–26.
——. 1988. "Learning, Price Effects and Income Growth." Working Paper, Princeton University, Princeton, N.J.
Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations* (1 ed.). London: W. Strahan.
Smith, Pamela, Nils B. Jostmann, Adam D. Galinsky, and Wilco W. van Dijk. 2008. "Lacking Power Impairs Executive Functions." *Psychological Science* 19 (5): 441–447.
Solow, Robert M. 1970. *Growth Theory—An Exposition*. Oxford: Oxford University Press.
——. 1962a. "Substitution and Fixed Proportions in the Theory of Capital." *Review of Economic Studies* 24 (June): 207–218.
——. 1962b. "Technical Progress, Capital Formation, and Economic Growth." In *Papers and Proceedings of the Seventy-Fourth Annual Meeting of the American Economic Association*, *American Economic Review* 52 (2): 76–86.
——. 1960. "Investment and Technical Progress." In *Mathematical Methods in the Social Sciences, 1959: Proceedings of the First Stanford Symposium*, ed. K. J. Arrow, S. Karlin, and P. Suppes, 89–104. Stanford, Calif.: Stanford University Press.

———. 1957. “Technical Change and the Aggregate Production Function.” *Review of Economics and Statistics* 39 (3): 312–320.
———. 1956. “A Contribution to the Theory of Economic Growth.” *Quarterly Journal of Economics* 70: 65–94.
Solow, Robert M., J. Tobin, M. E. Yarri, and C. C. von Weizacker. 1966. “Neo-Classical Growth with Fixed Factor Proportions.” *Review of Economic Studies* 33 (April): 79–116.
Spence, A. Michael. 1983. “Contestable Markets and the Theory of Industry Structure: A Review Article.” *Journal of Economic Literature* 21 (3, September): 981–990.
———. 1981. “The Learning Curve and Competition.” *Bell Journal of Economics* 12 (1): 49–70.
———. 1973. “Job Market Signaling.” *Quarterly Journal of Economics* 87 (3): 355–374.
Steele, Claude M. 2010. *Whistling Vivaldi and Other Clues to How Stereotypes Affect Us*. New York: Norton.
Stiglitz, Joseph E. 2014a. “Intellectual Property Rights, the Pool of Knowledge, and Innovation.” NBER Working Paper 20014.
———. 2014b. “Creative Destruction.” Working Paper, Columbia University.
———. 2014c. “Competition and Innovation.” Working Paper, Columbia University.
———. 2014d. “Unemployment and Innovation.” NBER Working Paper 20670.
———. 2014e. “Leaders and Followers: Perspectives on the Nordic Model and the Economics of Innovation.” NBER Working Paper 20493. Revised version forthcoming in 2015 in *Journal of Public Economics*.
———. 2014g. “Tapping the Brakes: Are Less Active Markets Safer and Better for the Economy?” Paper prepared for presentation at Atlanta Federal Reserve Conference, April 15, 2014.
———. 2013a. “Institutional Design for China’s Innovation System: Implications for Intellectual Property Rights.” In *Law and Economic Development with Chinese Characteristics: Institutions for the 21st Century*, eds. D. Kennedy and Joseph E. Stiglitz. New York: Oxford University Press.
———. 2013b. “Intellectual Property Rights, the Pool of Knowledge, and Innovation.” Working Paper, Columbia University.
———. 2012a. “Contagion, Liberalization, and the Optimal Structure of Globalization.” *Journal of Globalization and Development* 1 (2), article 2, 45 pages.
———. 2012b. *The Price of Inequality: How Today’s Divided Society Endangers Our Future*. New York: W. W. Norton.
———. 2012c. “Rebuttal Testimony of Dr. Joseph Stiglitz on Behalf of Respondents Responding to Direct Testimony of Dr. Stern.” In the Matter of Certain Microprocessors, Components Thereof, and Products Containing Same, United States International Trade Commission, Investigation No. 337-TA-781.

——. 2012d. “Direct Testimony of Joseph E. Stiglitz on Behalf of Respondents.” In the Matter of Certain Microprocessors, Components Thereof, and Products Containing Same, United States International Trade Commission, Investigation No. 337-TA-781.

——. 2011. “Macroeconomic Fluctuations, Inequality, and Human Development.” *Journal of Human Development and Capabilities* 13 (1): 31–58.

——. 2010a. “Evolutionary Theory and the Current Economic Crisis.” Presented at the American Economic Association annual meeting, January 2010. http://www2.gsb.columbia.edu/faculty/jstiglitz/download/papers/2010_Evolutionary_Theory.pdf.

——. 2010b. *Freefall: America, Free Markets, and the Sinking of the World Economy*. New York: W. W. Norton.

——. 2010c. Introduction to *Capitalism, Socialism and Democracy* by Joseph A. Schumpeter, ix–xiv. London: Rutledge.

——. 2010d. “Learning, Growth, and Development: A Lecture in Honor of Sir Partha Dasgupta.” Presented at the World Bank’s Annual Bank Conference on Development Economics 2010: Development Challenges in a Post-Crisis World. Published in French as “Apprentissage, croissance et développement: conférence en l’honneur de Sir Partha Dasgupta,” in *Revue D’Économie du Développement*, no. 4, December 2011, 19–86.

——. 2010e. “Risk and Global Economic Architecture: Why Full Financial Integration May Be Undesirable.” *American Economic Review* 100 (2, May): 388–392.

——. 2010f. “A Social Democratic Agenda for a More Dynamic Indian Economy: Creating an Innovative and Learning Society.” The 2010 Jawaharlal Nehru Memorial Lecture, New Delhi, November 18, available at http://www2.gsb.columbia.edu/faculty/jstiglitz/download/speeches/2010.11.18.Social.Democratic.Agenda.India.pdf.

——. 2010g. “Development-Oriented Tax Policy.” In *Taxation in Developing Countries*, ed. R. H. Gordon, 11–36. New York: Columbia University Press.

——. 2009. “Information and Economic Analysis.” In *Selected Works of Joseph E. Stiglitz*, vol. 1, *Information and Economic Analysis*, 29–52. Oxford: Oxford University Press.

——. 2008a. “Capital Market Liberalization, Globalization, and the IMF.” In *Capital Market Liberalization and Development*, ed. Joseph E. Stiglitz and J. A. Ocampo, 76–100. New York: Oxford University Press. (Earlier version published as “Capital-Market Liberalization, Globalization and the IMF.” *Oxford Review of Economic Policy* 20 [1, Spring 2004]: 57–71.)

——. 2008b. “The Economic Foundations of Intellectual Property.” *Duke Law Journal* 57 (6, April): 1693–1724. Sixth annual Frey Lecture in Intellectual Property, Duke University, February 16, 2007.

———. 2008c. "Is There a Post-Washington Consensus Consensus?" In *The Washington Consensus Reconsidered: Towards a New Global Governance*, ed. Narcis Serra and Joseph E. Stiglitz, 41–56. New York: Oxford University Press.

———. 2008d. "Toward a General Theory of Consumerism: Reflections on Keynes' *Economic Possibilities for Our Grandchildren*." In *Revisiting Keynes: Economic Possibilities for Our Grandchildren*, ed. G. Piga and L. Pecchi, 41–87. Cambridge, Mass.: MIT Press.

———. 2008e. "Regulating Multinational Corporations: Towards Principles of Cross-Border Legal Frameworks in a Globalized World Balancing Rights with Responsibilities." *American University International Law Review* 23(3): 451–558. Grotius Lecture presented at the 101st Annual Meeting of the American Society for International Law, Washington, D.C., March 28, 2007.

———. 2006a. *Making Globalization Work*. New York: Norton.

———. 2006b. "Samuelson and the Factor Bias of Technological Change." In *Samuelsonian Economics and the Twenty-First Century*, ed. M. Szenberg, L. Ramrattan, and A. Gottesman, 235–251. New York: Oxford University Press.

———. 2004. "Towards a Pro-Development and Balanced Intellectual Property Regime." Keynote address presented at the Ministerial Conference on Intellectual Property for Least Developed Countries, World Intellectual Property Organization (WIPO), Seoul, October 25.

———. 2003. *Roaring Nineties*. New York: Norton.

———. 2002a. *Globalization and Its Discontents*. New York: Norton.

———. 2002b. "Information and the Change in the Paradigm in Economics." Abbreviated version of Nobel lecture. *American Economic Review* 92 (3, June): 460–501.

———. 2001a. "Democratic Development as the Fruits of Labor." In *The Rebel Within*, ed. Ha-Joon Chang, 279–315. London: Wimbledon Publishing. (Originally the keynote address at the Industrial Relations Research Association, Boston, January 2000.) Shortened version: *Perspectives on Work* 4 (1): 31–38.

———. 2001b. "Participation and Development: Perspectives from the Comprehensive Development Paradigm." In *Democracy, Market Economics & Development: An Asian Perspective*, ed. Farrukh Iqbal and Jong-Il You, 49–72. Washington, D.C.: World Bank. Also in *The Rebel Within*, ed. Ha-Joon Chang, 220–249. London: Wimbledon.

———. 2001c. "Challenges in the Analysis of the Role of Institutions in Economic Development." In *Villa Borsig Workshop Series 2000: The Institutional Foundations of a Market Economy*, ed. Gudrun Kochendorfer-Lucius and Boris Pleskovic, 15–28. German Foundation for International Development (DSE).

———. 2000a. "Capital Market Liberalization, Economic Growth, and Instability." In *World Development* 28 (6): 1075–1086.

——. 2000b. "Formal and Informal Institutions." In *Social Capital: A Multifaceted Perspective*, ed. P. Dasgupta and I. Serageldin, 59–68. Washington, D.C.: World Bank.

——. 2000c. "Whither Reform? Ten Years of Transition." In *Annual World Bank Conference on Economic Development*, ed. B. Pleskovic and Joseph E. Stiglitz, 27–56. Washington, D.C.: World Bank.

——. 1999a. "Knowledge as a Global Public Good." In *Global Public Goods: International Cooperation in the 21st Century*, ed. Inge Kaul, Isabelle Grunberg, and Marc A. Stern, United Nations Development Programme, 308–325. New York: Oxford University Press.

——. 1999b. "Knowledge for Development: Economic Science, Economic Policy, and Economic Advice." In *Annual World Bank Conference on Development Economics*, ed. B. Pleskovic and Joseph Stiglitz, 9–58. Washington, D.C.: World Bank.

——. 1999c. "More Instruments and Broader Goals: Moving Toward the Post-Washington Consensus." In *Development Issues in the 21st Century*, ed. G. Kochendorfer-Lucius and B. Pleskovic, 11–39. Berlin: German Foundation for International Development. Also in *The Rebel Within*, ed. Ha-Joon Chang, 17–56. London: Wimbledon, 2001. Also in *WIDER Perspectives on Global Development*, ed. United Nations University—World Institute for Development Economics Research, 16–48. Houndmills, UK: Palgrave MacMillan, 2005. (Originally presented as the 1998 WIDER Annual Lecture, Helsinki, January 1998; also keynote address at Villa Borsig Winter Workshop, February 1998.)

——. 1999d. "Taxation, Public Policy and the Dynamics of Unemployment." *International Tax and Public Finance* 6: 239–262. (Paper presented to the Institute of International Finance, Cordoba, Argentina, August 24, 1998.)

——. 1998a. "An Agenda for Development in the Twenty-First Century." In *Annual World Bank Conference on Development Economics 1997*, ed. Joseph E. Stiglitz and B. Pleskovic, 17–31. Washington, D.C.: World Bank.

——. 1998b. "Pareto Efficient Taxation and Expenditure Policies, With Applications to the Taxation of Capital, Public Investment, and Externalities." Presented at conference in honor of Agnar Sandmo, Bergen, January.

——. 1998c. "Towards a New Paradigm for Development: Strategies, Policies and Processes." Ninth Raul Prebisch Lecture delivered at the Palais des Nations, Geneva, October 19, UNCTAD. Published as chapter 2 in *The Rebel Within*, ed. Ha-Joon Chang, 57–93. London: Wimbledon Publishing, 2001.

——. 1996. "Some Lessons from the East Asian Miracle." *World Bank Research Observer* 11 (2, August): 151–177.

——. 1995a. "Social Absorption Capability and Innovation." In *Social Capability and Long-Term Economic Growth*, ed. Bon Ho Koo and D. H. Perkins, 48–81. New York: St. Martin's.

———. 1995b. "The Theory of International Public Goods and the Architecture of International Organizations." Background Paper No. 7, Third Meeting, High Level Group on Development Strategy and Management of the Market Economy, United Nations University—World Institute for Development Economics Research, Helsinki, Finland, July 8–10.

———. 1994a. "Economic Growth Revisited." *Industrial and Corporate Change* 3 (1): 65–110.

———. 1994b. "Endogenous Growth and Cycles." In *Innovation in Technology, Industries, and Institutions*, ed. Y. Shionoya and M. Perlman, 121–156. Ann Arbor: University of Michigan Press.

———. 1994c. *Whither Socialism?* Cambridge, Mass.: MIT Press.

———. 1992. "Banks Versus Markets as Mechanisms for Allocating and Coordinating Investment." In *The Economics of Cooperation: East Asian Development and the Case for Pro-Market Intervention*, ed. J. A. Roumasset and S. Barr, 15–38. Boulder, Colo.: Westview. (Presented at Investment Coordination in the Pacific Century: Lessons from Theory and Practice Conference, University of Hawaii, January 1990.) Reprinted in *Selected Works of Joseph E. Stiglitz*, vol. 2, *Information and Economic Analysis: Applications to Capital, Labor, and Product Markets*, 258–272. Oxford: Oxford University Press, 2013.

———. 1990. "Comments: Some Retrospective Views on Growth Theory." In *Growth, Productivity, Unemployment: Essays to Celebrate Bob Solow's Birthday*, ed. Peter Diamond. Cambridge, Mass.: MIT Press.

———. 1989. "Monopolistic Competition and the Capital Market." In *The Economics of Imperfect Competition and Employment—Joan Robinson and Beyond*, ed. G. Feiwel, 485–507. New York: New York University Press; also published in *The Monopolistic Competition Revolution in Retrospect*, ed. S. Brakman and B. Heijdra, 49–67. Cambridge, UK: Cambridge University Press, 2004.

———. 1988. Economic Organization, Information, and Development." In *Handbook of Development Economics*, ed. H. Chenery and T. N. Srinivasan, 185–201. Amsterdam: Elsevier Science.

———. 1987a. "Learning to Learn, Localized Learning and Technological Progress." In *Economic Policy and Technological Performance*, ed. P. Dasgupta and P. Stoneman, 125–153. Cambridge, UK: Cambridge University Press.

———. 1987b. "On the Microeconomics of Technical Progress." In *Technology Generation in Latin American Manufacturing Industries*, ed. Jorge M. Katz, 56–77. London: Macmillan. (Originally presented to IDB-CEPAL Meetings, Buenos Aires, November 1978.)

———. 1987c. "Technological Change, Sunk Costs, and Competition." *Brookings Papers on Economic Activity* 3: 883–947.

———. 1986a. "Theory of Competition, Incentives and Risk." In *New Developments in the Analysis of Market Structure*, ed. Joseph E. Stiglitz and F. Mathewson, 399–449. Cambridge, Mass.: MIT Press.

——. 1986b. “Toward a More General Theory of Monopolistic Competition.” In *Prices, Competition, & Equilibrium*, ed. M. Peston and R. Quandt, 22–69. Oxford: Philip Allan/Barnes & Noble Books.

——. 1982a. “The Inefficiency of the Stock Market Equilibrium.” *Review of Economic Studies* 49 (2, April): 241–261.

——. 1982b. “Ownership, Control and Efficient Markets: Some Paradoxes in the Theory of Capital Markets.” In *Economic Regulation: Essays in Honor of James R. Nelson*, ed. Kenneth D. Boyer and William G. Shepherd, 311–341. Michigan State University Press. Reprinted in *The Selected Works of Joseph E. Stiglitz, Volume II: Information and Economic Analysis: Applications to Capital, Labor, and Product Markets*, 99–114. Oxford: Oxford University Press, 2013.

——. 1982c. “Information and Capital Markets.” In *Financial Economics: Essays in Honor of Paul Cootner*, ed. William F. Sharpe and Cathryn Cootner, 118–158. New Jersey: Prentice Hall. Also NBER Working Paper 678. Reprinted in *The Selected Works of Joseph E. Stiglitz, Volume II: Information and Economic Analysis: Applications to Capital, Labor, and Product Markets*, 55–84. Oxford: Oxford University Press, 2013.

——. 1981. “Potential Competition May Reduce Welfare.” *American Economic Review* 71 (2, May): 184–189.

——. 1977. “Monopoly, Non-Linear Pricing and Imperfect Information: The Insurance Market.” *Review of Economic Studies* 44 (3, October): 407–430.

——. 1975a. “Information and Economic Analysis.” In *Current Economic Problems: Proceedings of the Association of University Teachers of Economics, Manchester, 1974*, ed. J. M. Parkin and A. R. Nobay, 27–52. Cambridge, UK: Cambridge University Press.

——. 1975b. “The Theory of Screening, Education and the Distribution of Income.” *American Economic Review* 65 (3, June): 283–300. Reprinted in *Selected Works of Joseph E. Stiglitz, Volume I: Information and Economic Analysis*, 99–121. Oxford: Oxford University Press, 2009.

——. 1974a. “Theories of Discrimination and Economic Policy.” In *Patterns of Racial Discrimination*, ed. G. von Furstenberg et al., 5–26. Lexington, Mass.: D. C. Heath/Lexington.

——. 1974b. “Alternative Theories of Wage Determination and Unemployment in L.D.C.’s: The Labor Turnover Model,” *Quarterly Journal of Economics* 88 (2): 194–227. Subsequently published in *Development Economics*, *1*, ed. D. Lal, 288–321. Elgar, 1992. Reprinted in *The Selected Works of Joseph E. Stiglitz, Volume II: Information and Economic Analysis: Applications to Capital, Labor, and Product Markets*, 461–487. Oxford: Oxford University Press, 2013.

——. 1972. “On the Optimality of the Stock Market Allocation of Investment.” *Quarterly Journal of Economics* 86 (1, February): 25–60. (Presented to the Far Eastern Meetings of the Econometric Society, June 1970, Tokyo, Japan.)

——. 1969. "The Effects of Income, Wealth and Capital Gains Taxation on Risk-Taking." *Quarterly Journal of Economics* 83(2): 263–283.

——. 1967. "A Two-Sector, Two Class Model of Economic Growth." *Review of Economic Studies* 34 (April): 227–238.

Stiglitz, Joseph E., and José Antonio Ocampo, eds.. 2008. *Capital Market Liberalization and Development*. New York: Oxford University Press.

Stiglitz, Joseph E., and M. Uy. 1996. "Financial Markets, Public Policy, and the East Asian Miracle." *World Bank Research Observer* 11 (2, August): 249–276.

Stiglitz, Joseph E., and Hirofumi Uzawa, eds. 1969. *Readings in the Modern Theory of Economic Growth*. Cambridge, Mass.: MIT Press.

Stiglitz, Joseph E., and Scott J. Wallsten. 1999. "Public-Private Technology Partnerships: Promises and Pitfalls." *American Behavioral Scientist* 43 (1, September): 52–74. Also published in *Public-Private Policy Partnerships*, ed. P. Rosenau, 37–59. Cambridge, Mass.: MIT Press, 2000.

Stiglitz, Joseph E., and Andrew Weiss. 1981. "Credit Rationing in Markets with Imperfect Information." *American Economic Review* 71 (3, June): 393–410.

Stiglitz, Joseph E., José Antonio Ocampo, Shari Spiegel, Ricardo Ffrench-Davis, and Deepak Nayyar, eds. 2006. *Stability with Growth: Macroeconomics, Liberalization, and Development*. The Initiative for Policy Dialogue Series. Oxford: Oxford University Press.

Stiglitz, Joseph E., Amartya Sen, and Jean-Paul Fitoussi. 2010. *Mismeasuring Our Lives: Why GDP Doesn't Add Up*. New York: New Press. 2008 report available at http://www.stiglitz-sen-fitoussi.fr/en/index.htm.

Stokey, Nancy. 1986. "Dynamics of Industry Wide Learning." In *Economic Equilibrium Analysis: Essays in Honor of Kenneth Arrow*, Vol. II., ed. W. P. Heller, R. M. Starr, and David Starrett. Cambridge University Press.

Stolper, Wolfgang F., and Paul A. Samuelson. 1941. "Protection and Real Wages." *Review of Economic Studies* 9 (1): 58–73.

Sunstein, Cass. 2001. *Republic.com*. Princeton, N.J.: Princeton University Press.

Sutch, R. 2010. "The Unexpected Long-Run Impact of the Minimum Wage: An Educational Cascade." In P. W. Rhode, J. L. Rosenbloom, and D. Weiman (eds.), *Economic Evolution and Revolution in Historical Time*. Palo Alto, Calif.: Stanford University Press 387–418.

Sutton, John. 1991. *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*. Cambridge, Mass.: MIT Press.

Thaler, Richard H., and Cass R. Sunstein. 2008. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. New York: Penguin.

t'Hoen, Ellen F. M. 2003. "TRIPS, Pharmaceutical Patents and Access to Essential Medicines: Seattle, Doha and Beyond." In *Economics of AIDS and Access to HIV Care in Developing Countries: Issues and Challenges*, ed. J. P. Moatti, B. Coriat, Y. Souteyrand, et al., 39–67. Paris: ANRS.

Thompson, Peter. 2010. "Learning by Doing." In *Handbook of the Economics of Innovation*, Vol. 1, ed. Bronwyn H. Hall and Nathan Rosenberg, 429–476. Amsterdam: North-Holland.

Topalova, Petia. 2010. "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India." *American Economic Journal: Applied Economics* 2 (4): 1–41.

United Nations Development Programme (UNDP). 2003. *Making Global Trade Work for People*. See Malhotra.

Uzawa, Hirofumi. 1968. "The Penrose Effect and Optimum Growth." *Economic Studies Quarterly* 19: 1–14.

——. 1965. "Optimum Technical Change in an Aggregate Model of Economic Growth." *International Economic Review* 6 (1): 18–31.

Varian, Hal. 2011. "Federalism Offers Opportunities for Casual Experimentation." *Economist* April 27. http://www.economist.com/node/21256696.

Violante, Giovanni. No date. "Skill-Based Technical Change." Working Paper, New York University. http://www.econ.nyu.edu/user/violante/Books/sbtc_january16.pdf.

Vives, X. 2008. "Innovation and Competitive Pressure." *Journal of Industrial Economics* 56 (3): 419–469.

Wacziarg, Romain. 2001. "Measuring the Dynamic Gains from Trade." *World Bank Economic Review* 15: 393–429.

Wacziarg, Romain, and Karen Horn Welch. 2003. "Trade Liberalization and Growth: New Evidence." NBER Working Paper 10152, National Bureau of Economic Research, Cambridge, Mass.

Walters, Carl J., and C. S. Holling. 1990. "Large-Scale Management Experiments and Learning by Doing." *Ecology* 71 (6): 2060–2068.

Weber, Steven. 2005. *The Success of Open Source*. Cambridge, Mass.: Harvard University Press.

Weinberg, Bruce A. 2006. "A Model of Overconfidence." Manuscript, Ohio State University.

Wessel, Maxwell. 2012. "Big Companies Can't Innovate Halfway." Blog of the *Harvard Business Review*, October 4. http://blogs.hbr.org/cs/2012/10/big_companies_cant_innovate_halfway.html.

Williams, Heidi L. 2013. "Intellectual Property Rights and Innovation: Evidence from the Human Genome." *Journal of Political Economy* 121 (1): 1–27.

Williamson, John. 2008. "A Short History of the Washington Consensus." In *The Washington Consensus Reconsidered: Towards a New Global Governance*, ed. Narcis Serra and Joseph E. Stiglitz, 41–56. New York: Oxford University Press.

——. 1993. "Democracy and the 'Washington Consensus.'" *World Development* 21 (8, August): 1329–1336. http://www.sciencedirect.com/science/article/pii/0305750X9390046C.

———. 1990. "What Washington Means by Policy Reform." In *Latin American Adjustments: How Much Has Happened?*, ed. J. Williamson. Washington, D.C.: Institute for International Economics.

Winter, Sidney G. 1993. "Patents and Welfare." *Industrial and Corporate Change* 2 (2): 211–231.

World Bank. 2015. *World Development Report 2015: Mind, Society, and Behavior*. New York: World Bank and Oxford University Press.

———. 2005. *World Development Report 2005: A Better Investment Climate for Everyone*. New York: World Bank and Oxford University Press.

———. 2001. *World Development Report 2000–2001: Attacking Poverty*. New York: World Bank and Oxford University Press.

———. 1999. *World Development Report 1998–99: Knowledge for Development*. New York: Oxford University Press.

———. 1993. *The East Asian Miracle: Economic Growth and Public Policy*. World Bank Policy Research Report. Oxford: Oxford University Press.

Wright, G. 1986. *Old South, New South: Revolutions in the Southern Economy Since the Civil War*. New York: Basic Books.

Yeyati, Eduardo Levy, and Alejandro Micco. 2007. "Concentration and Foreign Penetration in Latin American Banking Sectors: Impact on Competition and Risk." *Journal of Banking & Finance* Elsevier, 31 (6, June): 1633–1647.

Young, Alwyn. 2003. "Gold into Base Metals: Productivity Growth in the People's Republic of China During the Reform Period." *Journal of Political Economy* 111 (6): 1220–1261.

———. 1993. "Invention and Bounded Learning by Doing." *Journal of Political Economy* 101: 443–472.

———. 1991. "Learning by Doing and the Dynamic Effects of International Trade." *Quarterly Journal of Economics* 106: 369–406.

Zhu, X. 2012. "Understanding China's Growth: Past, Present, and Future." *Journal of Economic Perspectives* 26 (4, Fall): 103–124.

Zimmerman, M. B. 1982. "Learning Effects and the Commercialization of New Energy Technologies: The Case of Nuclear Power." *Bell Journal of Economics* 13: 297–310.