

THE WORLD BANK

Research Observer

OXFORD
UNIVERSITY PRESS

30257

INSIDE

Mark Gersovitz
and Jeffrey S. Hammer
on Infectious Diseases
and Public Policy

•

Martín Rama
on Globalization
and the Labor Market

•

Will Martin
on Developing Countries'
Changing Role
in World Trade

•

Karla Hoff
on Historical Paths
to Institutional Development

Mohsin S. Khan and
Sunil Sharma
on IMF Conditionality and
Country Ownership

Stefan Koerberle
on Whether Policy-based
Lending Should Still
Involve Conditionality

ISSN 0257-3032
www.wbro.oupjournals.org

T H E W O R L D B A N K

Research Observer

EDITOR

Shantayanan Devarajan, World Bank

COEDITORS

Gershon Feder, World Bank

Shahid Yusuf, World Bank

EDITORIAL BOARD

Paul Collier, World Bank

Susan Collins, Georgetown University

Angus Deaton, Princeton University

Barry Eichengreen, University of California–Berkeley

Emmanuel Jimenez, World Bank

Danny M. Leipziger, World Bank

Howard Pack, University of Pennsylvania

Michael Walton, World Bank

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

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

The views and interpretations expressed in this journal are those of the authors and do not necessarily represent the views and policies of the World Bank or of its Executive Directors or the countries they represent. The World Bank does not guarantee the accuracy of data included in this publication and accepts no responsibility whatsoever for any consequences of their use. When maps are used, the boundaries, denominations, and other information do not imply on the part of the World Bank Group any judgment on the legal status of any territory or the endorsement or acceptance of such boundaries.

For more information, please visit the Web sites of the *Research Observer* at www.wbro.oupjournals.org, the World Bank at www.worldbank.org, and Oxford University Press at www.oup-usa.org.

Research Observer

Volume 18 • Number 2 • Fall 2003

| | |
|---|-----|
| Infectious Diseases, Public Policy, and the Marriage of Economics and Epidemiology Mark Gersovitz and Jeffrey S. Hammer | 129 |
| Globalization and the Labor Market Martín Rama | 159 |
| Developing Countries' Changing Participation in World Trade Will Martin | 187 |
| Paths of Institutional Development: A View from Economic History Karla Hoff | 205 |
| IMF Conditionality and Country Ownership of Adjustment Programs Mohsin S. Khan and Sunil Sharma | 227 |
| Should Policy-Based Lending Still Involve Conditionality? Stefan G. Koeberle | 249 |
| Cumulative Index, 2000–2003 | 275 |

Keep up-to-date with the latest contents of *The World Bank Research Observer*
by registering for our eTOC service at <http://www3.oup.co.uk/jnls/tocmail>.
This service is freely available to all, no subscription required.

The World Bank Research Observer (ISSN 0257-3032) is published twice a year by Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009 for The International Bank for Reconstruction and Development/THE WORLD BANK. Communications regarding original articles and editorial management should be addressed to The Editor, *The World Bank Research Observer*, The World Bank, 1818 H Street, NW, Washington, D.C. 20433, USA. E-mail: researchobserver@worldbank.org.

Oxford University Press is a department of the University of Oxford. It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide.

Subscriptions: Subscription is on a yearly basis. The annual rates are US\$35 (£27 in UK and Europe) for individuals; US\$90 (£62 in UK and Europe) for academic libraries; US\$112 (£75 in UK and Europe) for corporations. Single issues are available for US\$21 (£16 in UK and Europe) for individuals; US\$57 (£39 in UK and Europe) for academic libraries; US\$70 (£47 in UK and Europe) for corporations. All prices include postage. Individual rates are applicable only when a subscription is for individual use and are not available if delivery is made to a corporate address. *Readers with mailing addresses in non-OECD countries and in socialist economies in transition are eligible to receive complimentary subscriptions on request by writing to the UK address below.* All subscription requests, single issue and back issue orders, changes of address, and claims for missing issues should be sent to:

North America: Oxford University Press, Journals Customer Service, 2001 Evans Road, Cary, NC 27513-2009, USA. Toll-free in the USA and Canada: 800-852-7323, or 919-677-0977. Fax: 919-677-1714. E-mail: jnlorders@oup-usa.org.

Elsewhere: Oxford University Press, Journals Subscriptions Department, Great Clarendon Street, Oxford OX2 6DP, UK. Tel: +44 1865 353907. Fax: +44 1865 353485. E-mail: jnl.cust.serv@oup.co.uk.

Advertising: Helen Pearson, Oxford Journals Advertising, P.O. Box 347, Abingdon SO, OX14 1GJ, UK. Tel/Fax: +44 1235 201904. E-mail: helen@oxfordads.com.

Back Issues: The current plus two back volumes are available from Oxford University Press at the *North America* contact information listed above. Previous volumes can be obtained from the periodicals Service Company, 11 Main Street, Germantown, NY 12526 USA, phone: 1-518-537-4700.

Requests for permissions, reprints, and photocopies: All rights reserved; no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without either prior written permission of the publisher (Oxford University Press, Journals Rights and Permissions, Great Clarendon Street, Oxford OX2 6DP, UK; tel: +44 1865 354490; fax: +44 1865 353485) or a license permitting restricted copying issued in the USA by the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923 (fax: 978-750-4470), or in the UK by the Copyright Licensing Agency Ltd., 90 Tottenham Court Road, London W1P 9HE, UK. Reprints of individual articles are available only from the authors.

Copyright: Copyright © 2003 The International Bank for Reconstruction and Development/THE WORLD BANK. It is a condition of publication in the journal that authors grant an exclusive license to The International Bank for Reconstruction and Development/THE WORLD BANK. However, requests for permission to reprint material found in the journal should come to Oxford University Press. This ensures that requests from third parties to reproduce articles are handled efficiently and consistently and will also allow the article to be disseminated as widely as possible. As part of the license agreement, authors may use their own material in other publications provided that the journal is acknowledged as the original place of publication and Oxford University Press as the publisher.

Indexing and Abstracting: *The World Bank Research Observer* is indexed and/or abstracted by ABI/INFORM, CAB Abstracts, Current Contents/Social and Behavioral Sciences, Journal of Economic Literature/EconLit, PAIS International, RePEc (Research in Economic Papers), Social Sciences Citation Index, and Wilson Business Abstracts. The microform edition is available through ProQuest (formerly UMI), 300 North Zeeb Road, Ann Arbor, MI 48106, USA.

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

Postal Information: *The World Bank Research Observer* (ISSN 0257-3032) is published twice a year by Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009. Send address changes to *The World Bank Research Observer*, Journals Customer Service Department, Oxford University Press, 2001 Evans Road, Cary, NC 27513-2009.

Infectious Diseases, Public Policy, and the Marriage of Economics and Epidemiology

Mark Gersovitz • Jeffrey S. Hammer

The assumption of rational choice helps in understanding how people respond to infectious diseases. People maximize their well-being by choosing levels of prevention and therapy subject to the constraints they face. Objectives and constraints are numerous, necessitating tradeoffs. For example, this approach predicts how people respond to changes in the risk of infection and to the availability of diagnostic tests. The combination of individual rationality with epidemiological models of infection dynamics predicts whether individual choices about infectious disease prevention and therapies produce the best possible social outcomes. If not, individuals' choices generate rationales for government interventions to influence the levels of preventive and therapeutic activities. Optimal policy usually means accepting endemic infection, but at a level lowered by a coordinated package of interventions. Economics combined with epidemiology provides much qualitative guidance on the design of such packages, including immunization programs.

The recent marriage of economics and epidemiology, though long delayed, promises to be fruitful. Public health means public policy, never more so than with respect to infectious diseases. Economics has much to say about public policy, and epidemiology has much to say about the dynamics of infectious diseases. This article reports on the first progeny of this marriage—vigorous and precocious in what they say about public policy, but with potential still to fulfill.

Infectious disease has two modes of transmission: from person to person, as with HIV/AIDS, or by vectors, such as mosquitoes, as with malaria or yellow fever. Both modes are strongly influenced by how people choose to act both individually and collectively through their governments. People make choices about prevention and therapy from among the options available to them. An important influence on these

choices is the probability of becoming infected, which varies with the proportion of people who are infectious, among other factors. Choices about prevention and therapy in turn affect the aggregate infection rate, producing a system of simultaneous feedback. Classic models of infectious disease (Anderson and May 1991) have not, however, incorporated such endogenous behavioral mechanisms into their dynamics.

One way to do so is to postulate that individuals make rational choices, the traditional assumption of economics. Once individual behavior is introduced into the analysis, a natural question is whether individual behavior is the most desirable based on social criteria or the aggregation of the costs and benefits for all people of each individual's behavior. Answering this question is the bread and butter of economics, essentially assessing the validity in the particular circumstances of the health sector of Adam Smith's hypothesis of the Invisible Hand.

When individuals do not take into account the effects of their actions on others, either negative or positive, there are externalities and a justification for government intervention to align the private costs and benefits with the social costs and benefits of decisions. A major theme in the economics of public health is how and why these externalities occur and what can be done about them.

A number of economists have taken up the task of applying economics to public policy toward infectious diseases, beginning with the foundation block of individual rational behavior and working up to aggregate dynamics and the design of policy. An important concept is the optimal policy package, the optimal rules for coordinating preventive and therapeutic expenditures over the course of an epidemic. Although doctors and patients naturally group actions into prevention (avoiding infection) and therapy (mitigating infection), this distinction is not a good guide to the design of an economically efficient policy package.

On some economic criteria and for some infections, there is no distinction between prevention and therapy. For instance, when optimally chosen the two types of inputs may move together over the course of an epidemic, increasing or decreasing in step. Such synchronization need not occur, however, and reliance on prevention may decline as reliance on therapies rises or vice versa. The targeting of inputs is important in determining this pattern, for instance, whether treatment is directed at everyone or just the infected.

Both prevention and therapy likely involve externalities. In this regard, the important influence on the design of the policy package is whether the discrepancy between individual and social incentives to prevent disease is the same as that between individual and social incentives to adopt therapies. These discrepancies are equal for diseases from which people recover only to become susceptible again, as with curable sexually transmitted diseases, such as chlamydia and gonorrhea. In these cases it is optimal for government to subsidize preventive and therapeutic activities equally. For other diseases, the government may need to provide quite different financial incentives for people to undertake preventions and therapies. The

situation of vector-borne diseases is even more complicated. These diseases typically respond to many preventive actions. Some preventive actions share economic attributes of therapies, and others differ both from therapies and from other preventive actions.

Many of the examples of decisions about infectious diseases in this article come from the experience of poor countries and, in particular, from the experience with HIV/AIDS. First, infectious diseases are much more prevalent in poor countries both because their geoclimatic environments are more conducive to many of these diseases, especially those involving vectors and macroparasites, and because individuals and governments have taken many fewer of the actions that control infectious diseases in richer countries. Second, much of the work by economists on infectious diseases has been on HIV/AIDS in both rich and poor countries. HIV/AIDS, a threat everywhere, has become the most important health concern in many of the heavily infected countries of Africa. But the article also discusses the economics of diseases from which people recover and diseases spread by vectors and explores options for dealing with diseases, such as vaccines and curative medicines, which are unfortunately not available for the control of HIV.

Constraints, Objectives, and Strategic Behavior in Individual Choice

The rationality postulate supposes that individuals make the best choices, as they see them, from the options available to them subject to the information they have. As with all applications of this postulate, two things need to be specified for it to be useful in making predictions. First are the constraints that the individual faces: the options for action as determined by the disease environment, the information the individual has about the environment, and the costs of being sick and dealing with the disease through prevention and therapies. Second are the objectives of the individual, or how the individual ranks different choices among these options.

Constraints

The biology of a disease sets many of the constraints. Once infected, does an individual recover and become immune or susceptible again or suffer chronic ill health or death? How does the individual value these circumstances? The answers to these questions determine part of the costs of being infected. Can individuals know when they are infected and infectious, and when others are infectious, whether through symptoms or a medical test? Can therapy alter the course of the disease? If so, the costs of the disease include the costs of illness as mitigated by therapy plus the costs of the type of therapy chosen.

Now moving backward to the point at which a susceptible person is exposed, is the disease transmitted easily or with difficulty and how? What preventive actions can a person take to lessen the risk of exposure to infection, and how much do these preventive actions cost? Is a vaccine available? Is a vector such as a mosquito involved, and how can it be combated? To what extent is the risk of infection determined by the overall rate of infection and hence the stage of the epidemic? Are the infectious people who put a susceptible individual at risk anonymous, as in the case of influenza, or plausibly identifiable to the susceptible, as with a sexually transmitted disease?

Information also constrains how individuals respond to infectious diseases. Two types of information are relevant to individual decisions about preventive and therapeutic actions. The first type is general information, including information about the existence of a disease, how it is and is not transmitted, whether infectious individuals can be asymptomatic, and whether preventive, palliative, and curative options are available and at what cost. An important question is whether such general information is the real constraint on people's decisions. For a newly emerging disease, such as HIV/AIDS in the 1980s, gaining general information through research and disseminating it to the population as a whole are critical. But after the initial phase there may be much less scope for affecting behavior through general information, an important exception being young people if parents and other adults do not effectively communicate knowledge. If people already have this type of information, providing it over and over is unlikely to change how people behave. If individuals make rational choices, then to be effective interventions need to affect individuals' constraints.

The second type of information is specific to the individual. It includes information about whether the individual is infected or infectious, whether the people with whom the person consorts are infectious, and the risks in the immediate environment, such as disease vectors or the presence of macroparasites in soil and water. There may be no way to tell without a test whether someone is infectious. Of course, people may have useful information about whether they are themselves more or less likely to be infected and, especially in the case of sexually transmitted diseases, about which prospective contacts are more or less likely to be infected. Such information is better than nothing—sometimes much better—and can be used to condition choices. Better yet is a medical test. If a medical test exists, the decision to be tested raises many strategic questions for individuals that fit well within the rationality paradigm, as explored later in this article.

Objectives and Tradeoffs

A person's decisions affecting health are influenced by objectives as well as constraints. Though reflecting simple common sense, these objectives are often neglected in the formulation of health policies. That oversight is important because

knowledge of people's objectives helps predict their behavior and because most economists believe in consumer sovereignty, which means taking into account how people value their own circumstances. Individuals want to minimize the money costs of prevention and curative or other therapeutic activities insofar as they pay these costs. In addition, all the costs of prevention, including all the things that people forgo to avoid being infected, are also costs of a disease just as much as the pain, fear, loss of income and other opportunities, early death, and other direct costs incurred as a result of becoming infected.

People value many things in addition to health. They value physical intimacy with other human beings, including sexual relations, whether for pleasure or procreation. But physical intimacy promotes the spread of infection, whether sexually transmitted diseases or diseases of proximity such as influenza, tuberculosis, and leprosy. To the extent that individuals are altruistic, they also value the costs paid by people whom they might infect, such as their sexual partners. People may also want their sexual partners to believe that they are sexually faithful and they may want their partners to be sexually faithful independently of the risk of infection. People also value moving around freely, even in environments that expose them to the risk of infection by vector-borne and macroparasitic diseases. They prefer not to boil water, sleep under bednets, or restrict their outdoor activities. They want the benefits of going into forests even at the risk of malaria, yellow fever, and plague or of working on their farm even at the risk of schistosomiasis (bilharzia) or onchocerciasis.

Thus, many goals are in conflict, requiring tradeoffs. A basic tradeoff is between expenditures to augment health through prevention or therapy and expenditures on other goods and services that augment well-being in other ways. As people sacrifice successive units of expenditures on things other than health to increase their expenditure on health, their valuation of these successive sacrifices tends to rise, sometimes sharply. This rising marginal utility of what is sacrificed is a natural brake on how many resources people want to divert from other expenditures to health. Most people do not want to dedicate all their resources to improving their health.

All these considerations are irrelevant if individuals lack choices. For infectious diseases people can almost always pursue various preventive actions and they may have the option of therapies. For many diseases there is a continuum of increasing effort for prevention and therapy. People can wash their hands or boil water thoroughly, more thoroughly, even more thoroughly, and so on. Similarly, people can pursue and comply with therapeutic regimens to varying degrees. In many cases the effectiveness of successive equal increases in health inputs tends to diminish. This diminishing marginal product of inputs is another brake on how far people are prepared to increase health expenditures, complementing the increasing valuation of additional forgone units of expenditure on things other than health.

Rational Fatalism

An important determinant of choices about prevention is the infection rate among potential partners, casual contacts, and vectors, which in turn is likely to depend on the overall infection rate. In the case of AIDS, an incurable disease, models of rational choice have found that the number of partners a person chooses may not be negatively related to the infection rate of their potential partners, a sort of rational fatalism.

To illustrate this point, assume that people do not care about infecting their partners. For ease of argument, assume further that people get great utility from one (randomly chosen) partner and declining but positive utility from subsequent ones. All these partners are equally likely to be infected (and infectious). At moderate levels of overall HIV infection, people choose to have more than one partner but less than the number they would choose in the absence of the infection. By assumption, if the infection probability of all partners rises substantially, people do not cut back to less than one partner. Because their first partner is likely to infect them, more than one partner poses little additional risk. People no longer cut back the number of partners after the first and might even increase the number of partners back to the maximum they would have had were there no disease at all.

This conclusion is, of course, the outcome of a specially chosen example. Under different assumptions, such as an infection rate rising from negligible to moderate levels, people might react in the opposite way, cutting back on the number of partners. The general point is that rational people who weigh both a desire for more partners and the probability of infection may choose to increase the number of partners in response to an increase in the probability that any one of their potential partners is infected. Thus there is a rational explanation for fatalistic and seemingly irrational behavior (Philipson and Posner 1993, p. 49; Kremer 1996).

The notion of rational fatalism is relevant to choices about the intensity of all sorts of exposure, not just to HIV/AIDS. For example, each day that a person extends a trip into malarial forests might be analogous to an additional partner, presuming that after having been infected one can continue to be active, at least while incubating the disease. Rational fatalism will be less likely if a pathogen exhibits superinfectivity, that is, when even an infected person who has not yet recovered or been cured can be further infected through additional exposures (as in the case of many macroparasitic diseases, such as schistosomiasis) and if each round of superinfection does increasing additional harm.

Strategic Behavior

Together, many of the preceding considerations suggest that a distinction can be made between diseases that imply strategic behavior and those that do not. Strategic behavior arises when the contacts that put a person at risk are not anonymous and a

person can think about the information that prospective contacts have about their own infection status and their incentives to cooperate in providing such information. Issues arise of negotiation, retaliation, and incentives to tell the truth. Such interactions are conditioned by whether people are entirely self-interested or at least partially altruistic in caring about whether they infect others.

Once again, it is a question of objectives and constraints. What are the boundaries of the decisionmaking unit? Oneself? Oneself and a regular sexual partner in the case of sexually transmitted disease? Nearby neighbors in case of diseases transmitted through effluent discharge? Sexually transmitted diseases are the classic case of a strategic disease, and HIV/AIDS has put these concerns at center stage. But the practice in Japan of wearing face masks when infected by a common cold or flu virus is another example of people taking precautions to avoid infecting others, even people who are not in their immediate circle of family and friends.

Testing

There are good reasons to be tested for disease, even a noninfectious one such as cancer, but testing raises the most complex considerations when the response to a disease may be strategic, as with HIV/AIDS. Although it seems obvious that people generally want to remain uninfected and probably prefer not to infect their contacts, they may or may not want to learn their infection status, and if they do learn it they may use such knowledge for many purposes.

In the case of HIV/AIDS the decision on whether to test can be influenced by several factors: the probability that one or one's potential partners are infected; whether one cares only about oneself (egoist) or also about infecting others (altruist), something to which observed testing behavior speaks; whether one can know confidently the HIV status of one's potential partners; what one's partner might do on learning that one has been tested and perhaps, in addition, on learning one's test result, such as dissolving the partnership or taking punitive measures; the costs of being tested, including going to a test facility and coming back for test results, which may not be negligible;¹ available treatments if HIV-positive; whether legacy issues are a concern if one is infected, for instance, arrangements for one's children; whether one wants to have children and therefore unprotected sexual relations; and whether one is concerned about infecting children during pregnancy or breastfeeding.

The list is long and without clear implications. Important in thinking about tests is this question: What difference will the test result make to a person's actions afterward? Some scenarios for the case of HIV testing follow.

- People in sexual partnerships who can learn each other's infection status and who want to (and believe they can) be mutually monogamous may want to test so that they can have unprotected relations if both test negative.

- People in sexual partnerships who can learn each other's infection status may want to test so they can have protected relations or cease relations if one tests negative and the other positive, known as a discordant partnership.
- People who fear infection, are altruistic, and cannot confidently learn their partner's status may have little reason to test. If such people are HIV-negative, they take precautions to protect themselves. If HIV-positive, they take precautions to protect their partners. Such people therefore take precautions regardless of their status and, absent considerations that are not sexually strategic, such as the last four in the preceding list of motivations, have no behavior that would be conditioned on knowing their status.
- People who are egoists may increase their level of activity if they test positive because they have nothing to lose and do not care about the consequences for their partners (Philipson and Posner 1993). People viewed by moderate-risk people as too high risk to be acceptable partners who then test negative may make themselves eligible for partnerships that will sometimes infect them.

In these last cases, Philipson and Posner (1993, p. 84) argue, "testing may increase the incidence of AIDS rather than being sure to decrease it." Mechoulan (2003) shows that testing can worsen the epidemic and lower people's overall level of well-being (inclusive of their enjoyment of sexual activity) in a model of egoists who cannot find out each other's test results, a severely limiting assumption. A research priority is to understand possible outcomes when people can ask to learn the status of prospective partners and can condition their behavior on either their partner's test results or on their partner's refusal to provide results, a restriction on the potential behavior of HIV-positive egoists.

There is some statistical information on people's choices about being tested for infectious diseases involving strategic decisions, and on what they do after receiving their results. In the 1996 Tanzanian Demographic and Health Survey, 60 percent of men and 91 percent of women who reported having had a conventional sexually transmitted disease in the previous 12 months claimed to have informed their partner (Gersovitz 2002). Of the infected men, 80 percent tried to avoid infecting their partners, 15 percent did not try, and 6 percent believed their partner to be infected already. Of the infected women, 52 percent tried to avoid infecting their partners, 7 percent did not try, and 41 percent believed their partner to be infected already.² Thus, these respondents seem to be broadly altruistic in their admittedly self-reported behavior.

But the benefits of testing are very different for a conventional sexually transmitted disease than for HIV. Conventional sexually transmitted diseases are curable, and people, once cured, do not want to be reinfected by their regular partners. So rational behavior leads to possibly very different behavior in the cases of HIV and of conventional sexually transmitted disease.

Gersovitz (2002) reports some findings on HIV testing from Demographic and Health Surveys for three East African countries: Kenya 1998, Tanzania 1996, and Uganda 1995. A not insignificant number of people have been tested, from 17 percent of Kenyan men to 4 percent of Tanzanian women. In all three countries more men have been tested than women, consistent with a core group model of infection in which men infect their long-term partners after becoming infected by a small group of high-sexual-activity women (Anderson and May 1991; Over and Piot 1993). In this model of infection a negative result for a man is sufficient for the couple both to be HIV-negative. At least two-thirds of the Kenyan, Tanzanian, and Ugandan men and women who have not been tested say they would like to be, paralleling results of a survey by the Global Program on AIDS (Ingham 1995). By contrast, Fylkesnes and others (1999) report that only 1.7 percent of 4,812 randomly selected Zambian men chose to be tested and returned for their results, despite the fact that 37 percent said they were willing to be tested and counseled.

In addition to a representative random survey like the Demographic and Health Survey, there are small-scale studies of HIV testing and associated behavior based on samples of convenience that provide important hints about how people perceive their strategic situation, sometimes suggesting egoistical motivations for many. First, significant numbers of people in these studies who are tested do not return for their results, in contrast to the high percentages of respondents in the Demographic and Health Surveys who report wanting to be tested.

Furthermore, Cartoux and others (1998) report that women who tested HIV-positive were less likely to return for their test results than women who tested negative at 10 of 13 HIV testing centers in 9 poor countries.³ Perhaps these women can guess the result based on their past experiences or perhaps their current circumstances leave them less behavior to condition on test results relative to women who test negative.

Second, significant numbers of people who are tested and learn their results do not tell their partners, especially those who test HIV-positive (Ryder and others 1991; Van der Straten and others 1995; Ladner and others 1996; Grinstead and others 2001). Such behavior hints at an egoistical outlook, although the circumstances of HIV-positive women may mean that they have much less opportunity to inform their partners or much less purpose for doing so than those who are HIV-negative. For instance, they might know that their partners were their only possible source of infection.

Studies also find that people who receive HIV-positive test results do take precautions, primarily condom use or abstinence or fewer casual partners (Campbell and others 1997; Grésenguet and others 2002). Discordant partners provide an important window on whether behavior is egoistical or altruistic once infection has occurred, but the small size and potential sampling bias of some of these studies qualify any conclusions.⁴ Partners who learn that their partnership is discordant tend to adopt safer practices and to follow them more consistently than concordant partners, especially concordant negative partners (Kamenga and others 1991; Allen and others 1992; Serwadda and

others 1995; Guimarães and others 2001; Roth and others 2001). Furthermore, discordant partners tend to adopt greater precautions if the man is the HIV-negative partner (Kamenga and others 1991; Allen and others 1992; Serwadda and others 1995).

Finally, there is not much evidence on instances of violence or other punitive actions after the disclosure of test results (Maman and others 2000). Grinstead and others (2001) report that these instances were rare in their studies in Kenya, Tanzania, and Trinidad and Tobago, although Temmerman and others (1995) recount that 13 of 66 women in Nairobi suffered serious consequences following disclosures to their partners.

Aggregate Dynamics of an Epidemic, Associated Externalities, and Public Health

The aggregate dynamics of an epidemic are determined in part by biology, including the mechanisms of infection; in part by behavior at the individual level; and in part by the actions of governments. There is a lot going on, and it easily ends up a jumble. Because people as individuals do not influence the actions of governments, one way economists try to disentangle the roles of individuals and governments is to think first about the choices of individuals in the absence of governments and how these choices would aggregate. The next stage is to think about the choices that governments would make if they could hypothetically control all decisions directly in the interests of all individuals taken together, for example, by determining choices about the level of prevention and therapy for every individual without any discretion by these individuals.

If there is a discrepancy between what would happen in these two situations, that suggests the existence of externalities and a role for governments to influence the actions of individuals to transform the first situation into the second using interventions that are actually available to governments, such as financial incentives, rather than the hypothetical direct control. This section builds on the description of individual choices in the preceding section and looks first at government's constraints and objectives, then at the aggregate dynamics of infection, briefly at many different types of externalities and what governments can do to offset them, and finally in some detail at the design of immunization strategies.

Social Constraints and Objectives

Governments, like individuals, face constraints and have objectives. Much of what constrains individual behavior constrains government behavior as well, such as the biology of the disease and the associated preventive and therapeutic possibilities.

Important additional constraints on governments are the choices by individuals as conditioned by their own constraints and objectives.

There is also the issue of government objectives. Although the rationality postulate provides a basis for predicting individual behavior, it is also closely related to most economists' solution to another problem—how to evaluate what happens as an outsider or policy analyst and therefore how the government should make decisions. We started with the view that individuals choose what is best for them individually, so it is not a big stretch to assert that actions should be valued the way individuals value them.

In the context of infectious diseases the sovereignty of individual preferences means that an evaluation of an epidemic is not based solely on infection rates but takes into account that individuals also value engaging in risky activities that lead to infection. For instance, there is no point in evaluating an intervention to prevent sexually transmitted diseases solely in terms of the effect on the prevalence of the disease without recognizing that people value physical intimacy. If they did not, presumably, there would be no epidemic to begin with. The cost of a disease is the total diminution in people's well-being consequent on all the ramifications of the existence of the disease; no economist should think otherwise.

When individuals do not take into account the consequences of their actions for others, however, there are externalities. In this case individuals' actions taken together do not lead to the best overall outcome for them, opening a role for government in public health. Thus just because the government respects the sovereignty of individual preferences and therefore uses the same method to value the consequences to specific individuals as those individuals do, that does not mean that the world as it exists is the best possible world. With externalities, individuals are not fully incorporating all the social consequences of each decision when they choose among them.

The following dissection of the structure of externalities in public health shows how the characteristics of the infection determine the qualitative nature of policies. The nature and extent of externalities provide guidance on how to make a division between private and public efforts to control disease. Individuals balance the private costs and benefits as they perceive them, and governments adjust the perceived private costs and benefits to coincide with the social costs and benefits by using their power to tax and subsidize.

These considerations provide a method for analyzing such common concerns as the possibility that innovations in health technology could reduce overall well-being by indirectly promoting riskier behavior and infections, a case of immiserizing innovation. For instance, there may be a concern that new, partially effective drugs or vaccines could lower the costs to any individual of risky activity, thereby leading to more infections than without such behavior change. The discussion of testing has already raised this class of possibilities. From the perspective of the person making

the decision, the additional dangerous behavior is worthwhile given the individual benefits of the behavior and the diminished personal consequences made possible by the innovation. Concerns about immiserizing innovation, therefore, have validity only to the extent that the changed behavior has implications for people other than the person making the decision so that the externality is worsened in a way that cannot be offset by the government at sufficiently low cost.

A rationale for government intervention based on externalities does not depend on either myopia (a failure by individuals to understand or incorporate the future evolution of the epidemic into their decisions) or any other discrepancy between the way governments and individuals evaluate outcomes. Of course, if such discrepancies did exist, there could be rationales for government intervention in addition to externalities.⁵

Estimating the costs imposed on others by an infectious person is a difficult but unavoidable problem in establishing the extent of an externality. What is the monetary equivalent of the pain of illness or loss of life? Economists have some answers. One approach derives individuals' valuation of risk to health and life from the wage premium they require to do jobs that pose these kinds of risks (Viscusi 1992; Jones-Lee 1994; Pauly 1995).

Another approach avoids the explicit valuation of illness and death altogether when comparing public policies, using cost-effectiveness analysis instead to derive priorities among inputs. The costs of an input are divided by a given outcome, usually number of lives or life years saved or cases of disease prevented, yielding a result such as cost per life saved or disease prevented. Inputs with the lowest cost in these terms would then be selected. But this method can be severely misleading for a variety of reasons (Hammer 1993; Auld and others 2002). Here, the focus is on weaknesses resulting from ignoring the ways in which people affect and value health outcomes.

For one thing, cost-effectiveness can produce very unsatisfactory implicit valuations of life. Hammer (1993, p. 20) gives the following example:

Consider a situation in which two drugs are available to treat a particular disease. Drug 1 changes the probability of avoiding death from 0.2 to 0.3 and costs \$5 per treatment. Drug 2 changes the probability from 0.2 to 0.25 and costs \$2 per treatment. The cost per life saved by drug 1 is \$50 ($\$5/[0.3-0.2]$), while lives saved by drug 2 cost \$40 ($\$2/[0.25-0.20]$), making drug 2 more cost effective.

Most people probably would opt for drug 1, though, provided they are willing to pay more than \$60 to save their life. For any imputed value of life greater than \$60, the value of the increased probability of recovery outweighs the extra cost of the drug . . . Cost-effectiveness ratios, while seeming to avoid the contentious issue of deciding on a monetary value of life, merely disguise an implicit valuation that may not reflect people's preferences.

The reason cost-effectiveness leads to an unattractive choice in this example is that it implicitly and implausibly assumes that two doses of drug 2 are as effective as but cheaper than one dose of drug 1. Unfortunately, the biology of chemotherapy is unlikely to work this way.

Second, cost-effectiveness fails to consider externalities (or other market failures) in the health sector. It prescribes public spending without incorporating any criterion for distinguishing when expenditure should be in the private sector, with all the advantages of market-based incentives for efficiency, and when public involvement is needed to offset market inadequacies, such as externalities. It may be that the greatest cost-effectiveness is realized in cancer therapy. But if this therapy would be paid for by a rich private citizen on his or her own behalf, there is no obvious reason to raise tax dollars, with the associated administrative and incentive costs, to pay for such therapy through the government, and certainly no reason based on externalities.

Third, most applications of cost-effectiveness focus exclusively on benefits arising from improvements in health—diminutions in disease and death. But many health inputs may bring other benefits, such as the ability to farm land that was previously inaccessible because the risk of disease was too high.

Aggregate Dynamics

The marriage of models of rational behavior by individuals and the conventional aggregate models of epidemiology provides a basis for understanding the dynamics of the overall infection rate, the scope for externalities, and optimal public policy.

Diseases transmitted person to person. For diseases that spread from person to person, the simplest model of mathematical epidemiology determines the change in the number of new infections per period:

$$\text{New infections} = \alpha(\text{total susceptible people}) (\text{aggregate infection rate}) \quad (1)$$

The equation assumes a homogeneous population in which people contact each other randomly and with equal probability. The probability per contact of a susceptible person's meeting an infected (and infectious) person is the proportion of infected people in the population, the aggregate infection rate. In more complex situations, there may be many groups of people, each homogeneous within itself but differing in the probability with which members contact members of other groups, given that all people have one contact and each group has its own probability that a member is infected (the group's infection rate). In the homogeneous model of equation 1 the product of the aggregate infection rate and the total number of susceptible people gives the total number of susceptible people who meet an infectious person when every susceptible person has one contact. The factor α is an adjustment incorporating both the number of contacts per person per period and the inherent infectiousness

of an infected person and susceptibility of a susceptible person. Here again, in a non-homogeneous model, different groups may have different numbers of contacts per unit of time or inherent infectiousness or susceptibility.

In addition to the number of new infections a fully specified epidemiological model also determines the change in the total number of infected people as the difference between the number of new infections (equation 1) and the reduced number of infected people as they recover and become susceptible again, recover and become immune, or die. Other equations in the model determine the change in the numbers of people in other health statuses (susceptible and immune) and the change in the total population (births and deaths), and thereby determine the infection rate itself (Anderson and May 1991).

At the aggregate level the discussion of rational choice provides a justification for making the infection parameter (α in equation 1) a function of preventive measures and making the probabilities of recovery or death (in the other equations of a complete model) functions of therapeutic measures. The values of these health inputs (preventions and therapies) are the outcome of individual and government choices.

The chosen level of prevention is likely to depend on the risk of infection, providing scope for dynamic feedback in a model of choice-based epidemiology. Thus models of disease transmission that allow for endogenous behavior imply that α is a function of the infection rate (prevalence) rather than a constant as in classic epidemiology, bringing in the notion of the prevalence elasticity of prevention (Philipson and Posner 1993; Kremer 1996).⁶ A distinction between a purely epidemiological model and a choice-based one, therefore, is that in the epidemiological model the proportion of (remaining) susceptible people who become infected in each period always rises as the infection rate rises, whereas in the choice-based model it may or may not rise, depending on whether and how much people push the transmission parameter (α in equation 1) down as the infection rate rises.⁷ Philipson (2000) reviews evidence from the United States that indicates that people's preventive behavior responds to prevalence; there are no similar studies for poor countries. Choices about therapy also depend on the risk of infection and therefore on prevalence. For instance, when it is very likely that someone will become reinfected there may be little incentive to incur all the costs of a cure.

Vector transmission. Many infections are transmitted by vectors rather than directly from person to person. In these cases a person's risk of infection depends directly on the proportion of the vector that carries the infectious agent and only indirectly on the proportion of people who are infected and who may in turn infect the vectors. Consequently, even the simplest models of such diseases have more parameters than the corresponding models of diseases that are transmitted directly from person to person. These parameters represent the scope for choice by individuals and governments.

In situations of vector transmission, common parlance classifies several economically different types of inputs as prevention. One type affects the vector directly—for instance, spraying mosquitos that transmit malaria. A second type affects the probability that an uninfected person will be infected by an infected vector but does not prevent uninfected vectors from becoming infected by an infected person. These inputs break the human–vector cycle at only one point; an example is wearing waterproof footwear to avoid schistosomiasis.

A third class of inputs affects the probability that an uninfected vector will be infected by an infected person, but does not prevent an uninfected individual from becoming infected by an infected vector. An example is a program that encourages people not to void parasitic eggs into an environment where the eggs can mature. A fourth class of inputs affects both the probability that an infected person will infect an uninfected vector and the probability that an infected vector will infect an uninfected person, for example, wearing protective clothing and using bed nets to lower the chance of being bitten by a mosquito vector. In addition to these categories of prevention, there are various therapies for many vector-borne infections.

Simulations by Gersovitz and Hammer (forthcoming b) show that these different preventive and the therapeutic elements of an optimal package respond in different ways to the stage of the epidemic (aggregate rates of infection of humans and vectors and the ratio of vectors to humans) and to the underlying parameters of the model, especially the prices for the different inputs and the costs of illness. For instance, the optimal steady-state use of each input falls as its own price rises, but the response of one input to changes in the prices of other inputs varies, so that different types of inputs are not uniformly either complements or substitutes.

Furthermore, the paths of the inputs are sensitive to the choice of group targeted (the uninfected, the infected, or everyone) and to changes in how people value the sacrifice of other consumption as additional resources are transferred to the health sector. In one example, changing the targeting of a curative input from the infected to everyone reverses the direction of its path to the steady state so that use of the curative input rises rather than falls on the path to a lower steady-state level of infection from the original steady state without any inputs. Mass targeting of undiagnosed populations makes sense if it is impossible or very costly to diagnose the disease and if the costs of treatment, including side effects, are low. In the case of schistosomiasis, for example, some analysts (Prescott 1987) believe that mass undiagnosed treatment is desirable, so this type of targeting issue is a real concern.

The conclusions in Gersovitz and Hammer (forthcoming b) depend on simulations of models with parameters that have no empirical foundation and are therefore only suggestive of possibilities in the design of optimal responses to vector-borne infections. In particular, information is lacking on the relationships between the levels of prevention and therapy and their associated costs, on one hand, and the effectiveness of each level of these inputs in diminishing transmission or mitigating infections that

have occurred on the other hand. The epidemiological literature provides essentially no information on the functional forms and parameters of epidemiological processes that doubtless underlie the findings of Gersovitz and Hammer. Firmer conclusions must await a better empirical understanding of the disease processes under alternative input packages. By raising possibilities, these simulations may stimulate empirical work on these missing elements with the goal of more realistic modeling.

Externalities and the Design of Health Interventions

Externality is a central concept in formulating public health policy. If individuals are not compensated for benefits or charged for costs that they generate for others outside of family and friends, they usually do not take those consequences into account in making their decisions, leading to an externality. Economists commonly propose subsidizing actions that produce benefits for people other than the person taking the action and taxing actions that produce neglected costs. By inducing decisionmakers to take into account the effects on others, government interventions make people internalize the externality.

In principle, governments can subsidize or tax privately chosen levels of prevention and therapy to the extent that these activities affect the health of others. In practice, it may be difficult or impossible to promote certain preventive and therapeutic activities through price-based incentives. For some diseases there are inputs that are marketed, such as medicines, and inputs that involve an individuals' nonmarketed and unobservable actions, such as avoiding places and times that vectors are active or taking medicines as prescribed. Subsidy or tax interventions may be infeasible for nonmarketed and unobservable actions for which there are no transactions and therefore no prices.

For some diseases that have both marketed and nonmarketed preventive and therapeutic activities, government interventions may have to be targeted only to the marketed components. These price-based interventions will not usually achieve the best outcome. First, in some cases, even subsidizing the marketed component at 100 percent (making it free) may not imply a high enough subsidy of the total costs (marketed plus imputed nonmarketed). A greater than 100 percent subsidy is, of course, infeasible; for instance, paying people to take condoms will presumably run an infinite bill as anyone, whether a user or not, is happy to get paid to carry them away. Even when less than a 100 percent subsidy on the marketed component is in principle adequate to achieve the desired subsidy of total costs, individuals may substitute subsidized inputs for unsubsidized inputs to produce improved health, creating inefficiencies.

For other diseases that involve both marketed and nonmarketed inputs, health professionals may be able to follow up to ensure that patients use nonmarketed inputs, as in the case of directly observed treatment short-course (DOTS), for ensuring

that tuberculosis patients take their medication. This approach tries to enforce behavior by observing it, rather than relying on price-based incentives, and government spending on enforcement is justified by success in offsetting the externality.⁸ Imposing a quarantine has a similar motivation. In the case of an incurable sexually transmitted disease such as AIDS, there are very limited opportunities for government subsidies to prevent infection (on tests, treatment for curable sexually transmitted diseases that facilitate HIV infection, condoms, and clean needle programs, but not directly on safe sexual activity or clean needle use). Direct observation of behavior is essentially infeasible in most situations that spread sexually transmitted diseases, although the Thai government has implemented a campaign to enforce condom use in brothels.

As mentioned, the increasing marginal cost of successive sacrifices of nonhealth expenditures and declining marginal productivity of health inputs together often make it undesirable to take the necessary steps to eradicate a disease, even when eradication is technically possible. Although wiping out a disease is desirable in principle, people rarely find that eradication is desirable given its costs. The question becomes one of developing a policy package that is the best possible without eradicating the disease.⁹ The package involves intervening to offset externalities on an ongoing basis, accepting that the tradeoff of social costs and benefits leads to a situation in which the disease is endemic but prevalence is lower than it would be without optimal government intervention.

Diseases that involve transmission from vectors to humans and back to vectors fit the externality model well. For example, a farmer who is infected with schistosomiasis and voids one stage of the parasite into an irrigation canal where it can mature into another stage that infects other farmers probably does not incorporate these external costs fully into decisionmaking, if he even considers them at all. For this type of infectious disease the people put at risk are often neither friends nor family about whom the farmer cares most, and in any case the way the farmer provides the next link in the chain of infection is not transparent. The same situation would seem to prevail in the case of diseases transmitted directly from person to person but relatively anonymously, perhaps before infected people are even aware that they are infectious. Influenza, such as the one that killed millions of people in 1919, is a good example.

Sexually transmitted diseases seem to allow more scope for altruism because a sexual partner is rarely anonymous and is often included among people whose interests an infected person wants to take into account. Still, as the evidence suggests, many people may behave egoistically and may not take others' interests into account in important aspects of their decisions about HIV/AIDS and other sexually transmitted diseases. Finally, unlike infectious diseases of all sorts, cardiovascular disease, cancer, and some other illnesses have no externality because individuals do not put others at risk of infection, so there is no externality rationale for government intervention, although there may be other rationales, such as a failure in insurance markets (not discussed in this article).

For infectious diseases it is useful to distinguish two types of externalities, the pure infection externality and the pure prevention externality (Gersovitz and Hammer forthcoming a). The pure infection externality arises if individuals do not take into account how their becoming infected affects the risks to others of becoming infected. The pure prevention externality operates regardless of whether the individual making choices becomes infected. An example would be spraying insecticides or draining swamps, which lower others' risks of becoming infected regardless of whether the person who pays for the preventive actions becomes infected.

This distinction between externalities provides some qualitative guidance for government interventions. Consider a quintessential infection externality: The disease is transmitted from person to person; people are either infected and infectious or well but susceptible to infection; and once cured, a person is again susceptible. In offsetting the externality, it is equally desirable to prevent someone from becoming infected as to cure them: The goal is to keep more people out of the infectious group than would otherwise be the case. Consequently, prevention and cures should be subsidized at exactly the same rate, a good instance when the two types of inputs are identical from an economic viewpoint (Gersovitz and Hammer forthcoming a). This result is not general but depends on the structure of the disease as defined by what happens to an infected individual. Recovery to susceptibility, recovery to immunity, or death all imply a different distribution of government intervention between prevention and therapy.

Vector-transmitted diseases provide further complications in the design of an intervention package.¹⁰ Interventions are not restricted to activities that either prevent people from infecting others or that treat the infected. Some inputs affect the vector and its ability to infect people. In the case of vector control, pure prevention externalities arise to an extreme degree. The government may have to pay for the whole program of vector control because control at the individual level may be impractical or ineffective; the infiltration of vectors from outside a single individual's perimeter of control may be overwhelming. For instance, spraying in one's own compound may have little effect on the probability of getting malaria if one's neighbors do nothing. Without a method of coordinating large groups of people, the definition of a government, the result will be that no one bothers to spray.

Miguel and Kremer (2001) estimate the externalities associated with a curative input—deworming school children in Kenya. The positive externalities to some children of treating other infected children may be large enough to justify paying children to be dewormed over and above paying all the costs of treating them. Miguel and Kremer also emphasize the need to incorporate the consequences of externalities in estimating the benefits of treatment if the treatment group and the associated external benefits are large enough to improve the health of untreated individuals in the control group who live in the same community. In this case the simple difference between the outcomes of randomly selected treated and untreated individuals would

understate the benefits of treatment to treated individuals and would entirely neglect the indirect benefits to the untreated.

Although the general notion of externalities is straightforward, seemingly paradoxical results are possible, especially when people differ in number of contacts and probabilities of being infectious (nonhomogeneous groups). For instance, consider a disease that spreads in a population made up of two groups, one with a high exposure to infection and one with a low exposure. For concreteness, say that the infection is HIV/AIDS and that the first group consists of people with many sexual partners and the second of people with relatively few partners. From an individual's perspective there is no doubt that lowering the number of partners will lower the probability of becoming infected, other things equal. But what happens to the infection rate in the population as a whole?

Assume initially that people in the low sexual activity group (as measured by number of partners) are not sexually active. In this case the epidemic may spread rapidly among people in the high-sexual-activity group. Now consider an increase in the activity level of people in the low-activity group, perhaps to one contact. Some of these contacts will be with people in the high-activity group and some will result in infections of people in the low-activity group. But if the people in the low-activity group have only one contact, they will not infect anyone else. In the meantime, some of the people in the high-activity group who are not yet infected will be less likely to meet infected people in the high-activity group because some of their contacts will be with people in the low-activity group, so they will be less likely to be infected. In principle, this effect could be so strong as to extinguish the epidemic among people in the high-activity group and therefore among the population as a whole, although at a cost to people in the low-activity group who become infected during the transition.¹¹

This example illustrates a particularly strong positive externality of increased activity by people in the low-activity group. Or, to turn this lesson around, the observation of an increased prevalence among people in the high-activity group need not reflect an increase in risky behavior by people in this group. Instead, it may reflect decreased risky behavior by people in the low-activity group. Indeed, high-activity people may even be decreasing their risky behavior somewhat.

Spurred by this type of stark and simplified example, Kremer and Morcum (1998) looked for this effect in a more realistic model of HIV transmission calibrated to data from the United Kingdom. They estimate that 8 of 10 heterosexual people have sufficiently low sexual activity levels (as measured by number of partners) that an increase in their sexual activity would decrease the national prevalence level. Thus, the theoretical result may not simply be a curiosity.

Understanding this phenomenon may be important for interpreting the dynamics of an epidemic with different groups (Whitaker and Rentin 1992), but its policy implications are less clear. Few people would want to encourage people with low

sexual activity to take risks with their health for some notion of the social good. Nonetheless, such a finding would seem to suggest the need to find policies to lower infection rates among people in high-sexual-activity groups who may actually be made more vulnerable by campaigns that decrease activity among people in low-activity groups.

Immunization: A Surprisingly Controversial and Important Source of Externalities

Public immunization programs are often cited as an example of policies to overcome externalities, precisely because people do not take into account the benefits arising from their acquired inability to infect others. Stiglitz (2000, p. 134) presents the arguments for this position: "While the main beneficiary of a vaccination may be the individual protected, and there is a significant marginal cost of vaccinating an additional individual, the public health benefits from universal vaccination—the reduced incidence of the disease, possibly its eradication—are benefits from which no one can be excluded."

Recently, Francis (1997) and Philipson (2000) have disputed this conventional view. Thus Philipson (2000, pp. 1763–64) remarks that:

relying on standard arguments about the positive external effects of disease prevention, economists often [argue] for an active public role in the prevention of infectious diseases, such as AIDS (see for instance Stiglitz 1997, p. 15). However, economists have rarely attempted to explain patterns of disease occurrence or to evaluate public interventions in the context of a society with individuals who do the best they can given their constraints. Such recent analysis...has cast...doubt on the old textbook arguments by economists.

The recent analysis on which Philipson makes his case, however, does not really support the abandonment of Stiglitz's and others' intuition because it is based on very narrow assumptions.

Francis (1997) provides a formal example of a dynamic model of infection propagation in which there is no externality, or at least none that justifies government intervention. Francis's assumptions are deliberately chosen to provoke further research. He assumes that there are no births of uninfected nonimmune people, that no one recovers from the disease, and that no one dies of any cause.

Gersovitz (2003) gives the following explanation for Francis's conclusions. When there are no births or deaths and no one is vaccinated, infection spreads (as determined by equation 1), the proportion of susceptible individuals falls continuously and the proportion of infected individuals rises continuously. As the proportion of infected individuals rises, so do the risk of infection and the benefits of immunity through vaccination for the remaining susceptible individuals. Eventually the

proportion of the population that is infected (and therefore the probability of infection) reaches a threshold at which the value of being vaccinated just equals the cost of vaccination (which is constant). If it is worth vaccinating anyone it is now worth vaccinating every susceptible individual because the proportion of infected individuals (and the probability of infecting the remaining susceptible individuals) will otherwise continue to rise; it cannot fall because no new susceptible individuals are born and no infected individuals die. The only way to maintain the steady-state value of the infection rate is to vaccinate every susceptible individual once the threshold is reached. In the steady state the proportion of susceptible individuals is therefore zero.

The form of the rule for vaccination is therefore a threshold rule and is the same whether governments or private individuals decide on vaccination. Furthermore, the value of the threshold is the same, and in this sense there is no externality, at least not one that justifies intervention. It might be tempting to infer that government would want to set a lower threshold that is attained earlier. But why? Without government intervention, once the private threshold is reached, everyone gets vaccinated anyway, and if everyone is vaccinated when anyone is, there is no one left to benefit from earlier vaccination, which would keep down the aggregate infection rate, so there is no point to the government's inducing an earlier time for vaccinating everyone. Hence the values of the thresholds are the same.

Gersovitz (2003) also considers the case in which there are new births and people die, but only as a proportion of the people who are alive so that any group of people would die out asymptotically, a conventional assumption made for tractability in mathematical epidemiology. The assumption of births and deaths is the only change from Francis's model; infected individuals are still assumed never to recover. In this case there are people, the newly born, who would benefit from a lower threshold infection rate for vaccination that is reached earlier so that when they enter the world the infection rate is as low as possible. Their elders decide about vaccination based only on their own benefits and costs, ignoring the benefit to newborns, even their own, because they take the aggregate infection rate as external to their own decisions about vaccination. After all, each member of the elder cohort is only one of many people who influence the aggregate infection rate. The government, however, recognizes that if the threshold were lowered, newborns will benefit and that it can induce a lower threshold by subsidizing immunization. In addition, once the government's threshold infection rate is reached, not everyone who is susceptible is vaccinated.

An assertion that it is optimal to vaccinate everyone when the threshold is reached would involve a contradiction: If a rule of universal immunization were adopted, all new susceptible individuals would be vaccinated, thereby diluting the infected pool as a proportion of the total population and lowering the infection rate below the threshold. But this conclusion contradicts the notion of a threshold infection rate after which everyone is vaccinated. Therefore the government's optimal

strategy is to begin vaccinating a proportion of susceptible individuals less than one once a threshold infection rate has been reached and to keep vaccinating members of the susceptible pool inclusive of newborns in such a way that the proportions of susceptible, vaccinated, and infected people remain constant.

Without government subsidization of immunization to equalize the two thresholds, private individuals follow the same qualitative rule of a threshold, but the threshold is higher because it does not incorporate the impact on the aggregate probability of others becoming infected. In effect, births introduce heterogeneity into the population: Not everyone is alive at the same time, and newborns cannot take decisions simultaneously with those born before them. The result of a positive subsidy obtains whether population growth is positive or not; all that is needed is that the birth rate be positive, not that it exceed the death rate.

Gersovitz (2003) also shows that if infected individuals recover and become susceptible again, recovery can play the same role as a positive birth rate in generating an externality and a motive for government intervention. For both births and recoveries there is a straightforward formula for the optimal subsidy of the cost of vaccination to internalize the externality. The absolute value of the subsidy depends negatively on the cost of the vaccination, the interest rate, and the natural death rate of the population and positively on the cost of infection, the birth rate of the population, and the recovery rate, but not at all on α , the transmission parameter of equation 1.

The purpose of government intervention is to improve welfare by lowering endemic prevalence. Only if the cost of infection is infinite or the cost of immunization is zero does the government's optimal threshold for beginning immunization equal zero so that the disease is eradicated (but only asymptotically because the population of infected individuals only dies out asymptotically, a modeling artifact of the assumption made for mathematical tractability that a constant fraction of the population dies in every period). In this case of an infinite cost of the disease or a zero cost of immunization, the government threshold coincides with the private threshold, and because both thresholds are set to zero there is no need for government intervention to bring about eradication.

Contrast these results on subsidization and eradication with those of Geoffard and Philipson (1997, p. 222), who believe that the inability to achieve eradication "brings into doubt classic justifications of Pigouvian subsidies aimed at solving the underprovision of vaccines due to their positive external effects."¹² A policy of eradication is not synonymous with a policy that optimally offsets the externality by using subsidies. Although eradication may well be too expensive relative to its benefits, there is still a rationale for the optimal subsidization of immunization to lessen the endemic prevalence of disease. Indeed, with current models dependent on the assumption that disease can only die out asymptotically, a research priority is modeling immunization when eradication is feasible in finite time.¹³

At this stage, it seems that the Francis-Philipson position against an externality-based rationale for vaccination does not extend beyond one generally inapplicable example, really more of a curiosity and a stimulus to further research than a guide to policy. More work, however, is needed to delineate the full scope of externalities associated with vaccinations.

Conclusion

This article has examined the implications of rational choice for both individual behavior and public policy. It argues that this way of looking at how people make decisions that affect their health helps in understanding what people do and what governments can and should do to change health outcomes. The structure of objectives and constraints is critical in building this framework. Biology and epidemiology are only two components of this framework but are important ones for determining constraints.

At the individual level models of rational choice generate hypotheses about what people do about their health, especially how they respond to increased risk of infection. This approach provides insight into rational fatalism, cases in which individuals react to increases in the probability of infection by decreasing certain types of preventive behavior. It also yields a typology of the consequences of the availability of tests based on the circumstances of individuals.

The greatest relevance of these models for public policy is for explicating the aggregate consequences of people's individual choices. Here the concept of externalities comes to the fore. Models of infectious diseases are rife with externalities, and the conclusion is that choices by individuals will almost never lead to the best outcome based on social criteria, even when the social criteria respect the sovereignty of individual preferences.

But the literature on externalities and infectious diseases does not stop with the general statement that private choices do not produce socially optimal outcomes. It provides guidance on how the government should intervene, usually through subsidies, to improve people's well-being as measured from a social perspective. It provides qualitative results on how interventions should affect incentives for prevention relative to therapy and, when there are many different types of prevention, on how to adjust the incentives to undertake them. It provides results on when and how to subsidize vaccinations as a function of such underlying parameters as the costs of illness and vaccination, the birth rate of susceptible individuals, and the recovery rate of infected individuals. These and other conclusions are a consequence of marrying economics and epidemiology.

For this marriage to be as productive as possible, however, more will be needed than simply a combination of what each side has already brought to the union.

Full consummation of the union requires quantitative information on the response of health outcomes to various preventive and therapeutic inputs. It also requires information on the choices individuals make when faced with different incentives. Efforts by both parties to establish this information should prove very fruitful.

Notes

Mark Gersovitz is Professor of Economics at the Johns Hopkins University. Jeffrey S. Hammer is Lead Economist, Development Research Group at the World Bank. Parts of this article draw on Gersovitz's (1999) Joseph Fisher Lecture at the University of Adelaide, in turn the basis for Gersovitz (2000).

1. For instance, when an HIV/AIDS testing and counseling center in Kampala, Uganda, replaced tests that required clients to return in two weeks with same-day results, the share of people tested who learned their test results rose from 79 percent to 100 percent (Kassler and others 1998). On the other hand, it may be that the additional 21 percent who learned their results had little or no use for the knowledge and that costs were not a deterrent.

2. Petry and Kingu (1996, p. 267–268) report that almost all people in their sample with a conventional sexually transmitted disease were aware of it, although only 25 percent of patients who reported sexually transmitted disease–like symptoms had a conventional sexually transmitted disease.

3. Among more recent studies with relatively large samples, Kawichai and others (2002) report that HIV-positive men were less likely and HIV-positive women equally likely to return for their tests as the corresponding HIV-negative groups. Grésenguet and others (2002) report no difference in serostatus between people who do and do not return for test results.

4. The ratio of male HIV-positive to female HIV-positive discordant partnerships is close to one in studies that report the serostatus of discordant couples (Gersovitz 1999). These results include: 69 male HIV-positive and 80 female HIV-positive in Kamenga and others (1991); 12 and 16 in Van der Straten and others (1995); 44 and 35 in Serwadda and others (1995); approximately 350 of each in McKenna and others (1997); 18 and 39 in Grinstead and others (2001); 43 and 23 in Roth and others (2001) and 22 and 21 in Hugonnet and others (2002). Carpenter and others (1999) present evidence that the probability of seroconversion of women in discordant partnerships is twice that of men. Even taking account of this finding, these ratios of male HIV-positive to female HIV-positive discordant partnerships suggest that a core-group model in which men become infected by very high-activity women and then infect their regular partners cannot be the whole story; too many women in the discordant regular partnerships are the infected partner.

5. Clear isolation of the dynamics of disease control and of externalities requires careful attention to the specification of private behavior. Thus Kremer (1996) posits that people care only about the chance of ever being infected and not about when they are infected. Such a utility function may be a useful first approximation, but it has clear limitations in conflicting with the criterion of present-discounted valuation of costs and benefits and therefore provides a poor basis for trading off expenditures for the control of epidemics against their (often future) benefits. Similarly, Geoffard and Philipson (1997) specify a demand function for vaccines, but only on the assumption that people choose to be vaccinated at birth or not at all and there is no discounting.

6. Most economists would find the notion of prevalence-elastic choices about prevention to be second nature. In a model of HIV/AIDS in a closed population with neither births nor deaths, however, Geoffard and Philipson (1996, p. 604) push the consequences of prevalence elasticity further than seems justified: "A growing prevalence and the public subsidy *compete* to induce protective activity, and this makes the timing of the public intervention a crucial factor in determining its economic efficiency. If the subsidy is not fast enough, the prevalence has made it irrelevant in inducing protection" (emphasis in original). This statement seems to mean that if government intervention has been delayed

sufficiently, there is no point in doing anything. But as argued later, however, externalities justify a government subsidy to induce prevention and achieve a lower rate of infection than would private behavior alone. Even if the government has waited and the infection rate has gone beyond what it should have been limited to, this situation should be only temporary. With newly born (or newly sexually active) people entering the pool of susceptible people all the time, the government can induce additional prevention and get the infection rate back down to the socially optimal level. As in many situations in life, better late than never.

7. In the case of the HIV/AIDS epidemic in Africa, traditional epidemiological models (for example, United Nations and World Health Organization 1991) have tended to underpredict the epidemic, presenting a major puzzle for understanding the dynamics of this epidemic. The possibility of a high prevalence elasticity—stressed by Philipson and Posner (1993) but omitted from traditional epidemiological models—and the consequent fall in α and the proportion of susceptible people who become infected as the epidemic evolves is not the solution. Indeed to the extent that highly prevalence-elastic behavior is operative, the puzzle is all the larger. Even if rational fatalism were operative, the actual α would be higher relative to the epidemiologists' (constant) estimate only if α were estimated from situations when infection rates were already fairly high but before rational fatalism became operative. If instead α were estimated from the incipient stages of the epidemic, rational fatalism could never lead to an actual α higher than the initial estimate because even rationally fatalistic people would not want to engage in more sexual activity than when there was almost no risk. In models of rational fatalism the only benefit people obtain from sexual activity is the activity itself, which is traded off against the (undesirable) risk of becoming infected. Within the confines of these models, the only people who would want to have more partners in the presence than in the absence of the disease would be people who wanted to engage in more risky behavior precisely because they valued the chance of becoming infected. A different set of considerations arises if one steps outside these models and thinks about people for whom sexual activity is part of relationships and family formation that imply fewer partners than people would otherwise prefer. If the epidemic were somehow to disrupt family formation with its associated restraint on the number of partners, then it is possible that sexual activity would increase as the epidemic progresses, but such speculation is outside the currently available models.

8. A person who gets cured provides a benefit to others whom he or she might have infected, but a person who gets cured in such a way as to induce a disease's resistance to drugs has a negative effect on others who become infected by making it harder for them to be cured. Traditional medical practice is to keep prescribing the best among a group of antibiotics until resistance to it arises and then to move to the next best and so on, "best" being determined by some combination of cost and efficacy. A number of economists have argued, however, that more than one antibiotic should be used simultaneously on the patient population (although not on any individual patient) and that doing so will minimize the costs of resistance (Laxminarayan and Brown 2001; Laxminarayan and Weitzman 2002; Laxminarayan 2003). By contrast, Salant (2003) emphasizes the difficulty of treating similar patients differently in the name of the social good and thereby questions the practicality of a multidrug strategy of minimizing drug resistance. Nonetheless, the negative externality associated with resistance would seem to justify unambiguously a public subsidy of programs, such as DOTS to ensure that people comply with drug regimens even under the traditional one-drug-at-a-time strategy, because noncompliance is a source of drug resistance.

9. Much of the literature that deals with the optimal use of inputs in mathematical epidemiology formulates the problem in such a way that inputs are used to the fullest extent or not at all. The reason is that this literature ignores the decreasing marginal product of inputs and the increasing marginal utility of the diverted resources that are necessary to pay for the inputs. Wickwire (1977) surveys early work; Sethi and Staats (1978), Greenhalgh (1988), and Hocking (1991) are more recent. The economists Geoffard and Philipson (1996, 1997) are similarly concerned primarily with the question of whether governments can eradicate a disease and not with whether it is optimal to do so or to choose instead among endemic rates of infection. The only mathematical discussion of when eradication dominates chronic control of an infection (and vice versa) that we know is by the economists Goldman and Lightwood (2002). Their analysis is of a somewhat special case because it only looks at

therapies, ignores the increasing marginal utility of diverted resources and uses a model in which the disease can only die out asymptotically. Nonetheless, it should be of pivotal importance to understanding this subject.

10. In the case of schistosomiasis, Wiemer (1987) looks at more than one input at a time, but the specification of costs and benefits implies that one of two inputs is used either to the maximum extent possible or not at all, a rather special case, and there is no discussion of externalities and the division of decisions between the private and public sectors.

11. Kremer (1996) shows that there can be multiple equilibria when people choose their levels of sexual activity endogenously and differ in their preferences about activity and the risk of infection.

12. Note that in the model of Geoffard and Philipson (1997) death is proportional to the population so that any group in the population never dies out in finite time. Consequently, eradication in their model is at best asymptotic. Furthermore, they do not provide a discussion of the social optimality of eradication under their assumption of heterogeneous demand for vaccines, so it is hard to contrast their findings directly with Gersovitz (2003). Finally, in their brief discussion of eradication, when the disease presumably may die out in finite time, Geoffard and Philipson (1997, p. 227) do not analyze when eradication is optimal but only provide weak upper and lower bounds on the costs and benefits of eradication. (See also the discussion in note 6.)

13. An overlapping generations (OLG) model is one way to develop this strategy; Mechoulan (2003) uses this approach in his analysis of testing.

References

- Allen, S., J. Tice, P. Van de Perre, A. Serufilira, E. Hudes, F. Nsengumuremyi, J. Bogaerts, C. Lindan, and S. Hulley. 1992. "Effect of Serotesting with Counselling on Condom Use and Seroconversion among HIV Discordant Couples in Africa." *British Medical Journal* 304:1605–9.
- Anderson, R. M., and R. M. May. 1991. *Infectious Diseases of Humans*. Oxford: Oxford University Press.
- Auld, C. M., C. Donaldson, C. Mitton, and P. Shackley. 2002. "Health Economics and Public Health." In R. Detels, J. McEwen, R. Beaglehole, and H. T. Tamaka, eds., *Oxford Textbook of Public Health*, 4th ed. Oxford: Oxford University Press.
- Campbell, C. H., M. E. Marum, M. G. Alwano-Edyegu, B. Dillon, M. Moore, and E. Gumisiriza. 1997. "The Role of HIV Counseling and Testing in the Developing World." *AIDS Education and Prevention* 9:92–104.
- Carpenter, L. M., A. Kamali, A. Ruberantwari, S. S. Malamba, and J. A. Whitworth. 1999. "Rates of HIV-1 Transmission within Marriage in Rural Uganda in Relation to the HIV Sero-Status of the Partners." *AIDS* 13:1083–89.
- Cartoux, M., N. Meda, P. Van de Perre, M. L. Newell, I. de Vincenzi, and F. Dabis. 1998. "Acceptability of Voluntary HIV Testing by Pregnant Women in Developing Countries: An International Survey." *AIDS* 12:2489–93.
- Francis, P. J. 1997. "Dynamic Epidemiology and the Market for Vaccinations." *Journal of Public Economics* 63:383–406.
- Fylkesnes, K., A. Haworth, C. Rosenvard, and P. Kwapa. 1999. "HIV Counselling and Testing: Over-emphasizing High Acceptance Rates a Threat to Confidentiality and the Right Not to Know." *AIDS* 13:2469–74.
- Geoffard, P.-Y., and T. Philipson. 1996. "Rational Epidemics and Their Public Control." *International Economic Review* 37:603–24.
- . 1997. "Disease Eradication: Private versus Public Vaccination." *American Economic Review* 87:221–30.

- Gersovitz, M. 1999. *Human Behaviour and the Transmission of Infectious Disease: An Economist's Perspective*. The Forty-Seventh Joseph Fisher Lecture in Commerce. Adelaide: University of Adelaide.
- . 2000. "A Preface to the Economic Analysis of Disease Transmission." *Australian Economic Papers* 39:68–83.
- . 2002. "The African HIV Epidemic as Seen through the Demographic and Health Surveys." Mimeo, The Johns Hopkins University, Department of Economics, Baltimore, Md.
- . 2003. "Births, Recoveries, Vaccinations and Externalities." In R. J. Arnott, B. Greenwald, R. Kanbur, and B. Nalebuff, eds., *Economics for an Imperfect World: Essays in Honor of Joseph Stiglitz*. Cambridge, Mass.: MIT Press.
- Gersovitz, M., and J. S. Hammer. Forthcoming a. "The Economical Control of Infectious Diseases." *Economic Journal*.
- . Forthcoming b. "Tax/Subsidy Policies toward Vector-Borne Infectious Diseases." *Journal of Public Economics*.
- Goldman, S. M., and J. Lightwood. 2002. "Cost Optimization in the SIS Model of Infectious Disease with Treatment." *Topics in Economic Analysis & Policy* 2(1).
- Greenhalgh, D. 1988. "Some Results on Optimal Control Applied to Epidemics." *Mathematical Biosciences* 88:125–58.
- Grésenguet, G., J. Schonou, B. Bassirou, J. D. Longo, J. E. Malkin, T. Brogan, and J. Belec. 2002. "Voluntary HIV Counseling and Testing: Experience among the Sexually Active Population in Bangui, Central African Republic." *Journal of Acquired Immune Deficiency Syndromes* 31:106–14.
- Grinstead, O. A., S. E. Gregorich, K. H. Choi, and T. Coates. 2001. "Positive and Negative Life Events after Counselling and Testing: The Voluntary HIV-1 Counselling and Testing Efficacy Study." *AIDS* 15:1045–52.
- Guimarães, M. D. C., C. Boschi-Pinto, and E. A. Castilho. 2001. "Safe Sexual Behaviour among Female Partners of HIV-Infected Men in Rio de Janeiro, Brazil." *International Journal of STD & AIDS* 12:334–41.
- Hammer, J. S. 1993. "The Economics of Malaria Control." *World Bank Research Observer* 8:1–22.
- Hocking, L. M. 1991. *Optimal Control: An Introduction to the Theory with Applications*. Oxford: Clarendon Press.
- Hugonnet, S., F. Moshia, J. Todd, K. Mugeye, A. Klokke, L. Ndeki, D. Ross, H. Grosskurth, and R. Hayes. 2002. "Incidence of HIV Infection in Stable Sexual Partnerships: A Retrospective Cohort Study of 1802 Couples in Mwanza Region, Tanzania." *Journal of Acquired Immune Deficiencies Syndromes* 30:73–80.
- Ingham, R. 1995. "AIDS: Knowledge, Awareness, and Attitudes." In J. Leland and B. Ferry, eds., *Sexual Behaviour and AIDS in the Developing World*. London: Taylor and Francis.
- Jones-Lee, M. W. 1994. "Safety and the Saving of Life: The Economics of Safety and Physical Risk." In R. Layard and S. Glaister, eds., *Cost-Benefit Analysis*. Cambridge: Cambridge University Press.
- Kamenga, M., R. Ryder, M. Jingu, and others. 1991. "Evidence of Marked Sexual Behavior Change Associated with Low HIV-1 Seroconversion in 149 Married Couples with Discordant HIV- Serostatus: Experience at an HIV Counselling Center in Zaire." *AIDS* 5:61–67.
- Kassler, W. J., M. G. Alwano-Edyegu, E. Marum, B. Biryahwaho, P. Kataaha, and B. Dillon. 1998. "Rapid HIV Testing with Same-Day Results: A Field Trial in Uganda." *International Journal of STD & AIDS* 9:134–38.
- Kawichai, S., D. D. Celentano, R. Chaifongsri, K. E. Nelson, K. Srithanaviboonchai, C. Natpratan, and others. 2002. "Profiles of HIV Voluntary Counseling and Testing of Clients at a District Hospital, Chiang Mai Province, Northern Thailand, from 1995 to 1999." *Journal of Acquired Immune Deficiencies Syndromes* 30:493–502.
- Kremer, M., 1996. "Integrating Behavioral Choice into Epidemiological Models of AIDS." *Quarterly Journal of Economics* 111:549–73.

- Kremer, M., and C. Morcum. 1998. "The Effect of Changing Sexual Activity on HIV Prevalence." *Mathematical Biosciences* 151:99–122.
- Ladner, J., V. Leroy, P. Msellati, M. Nyiraziraje, A. De Clercq, P. Van de Perre, and F. Dabis. 1996. "A Cohort Study of Factors Associated with Failure to Return for HIV Post-Test Counselling in Pregnant Women: Kigali, Rwanda, 1992–1993." *AIDS* 10:69–75.
- Laxminarayan, R., ed.. 2003. *Battling Resistance to Antibiotics and Pesticides*. Washington, D.C.: Resources for the Future.
- Laxminarayan, R., and G. M. Brown. 2001. "Economics of Antibiotic Resistance: A Theory of Optimal Use." *Journal of Environmental Economics and Management* 42:183–206.
- Laxminarayan, R., and M. L. Weitzman. 2002. "On the Implications of Endogenous Resistance to Medications." *Journal of Health Economics* 21:709–18.
- Maman, S., J. Campbell, M. D. Sweat, and A. C. Gielen. 2000. "The Interactions of HIV and Violence: Directions for Future Research and Interventions." *Social Science and Medicine* 50:459–78.
- McKenna, S. L., G. K. Muyinda, D. Roth, M. Mwali, N. N'gandu, A. Myrick, F. H. Priddy, V. M. Hall, A. A. von Lieven, J. R. Sabatino, and S. Allen. 1997. "Rapid HIV Testing and Counseling for Voluntary Testing Centers in Africa." *AIDS* 11:S103–10.
- Mechoulan, S. 2003. "HIV Testing: A Trojan Horse." Mimeo, University of Toronto, Department of Economics.
- Miguel, E., and M. Kremer. 2001. *Worms: Education and Health Externalities in Kenya*. NBER Working Paper 8481. National Bureau of Economic Research, Cambridge, Mass.
- Over, M., and P. Piot. 1993. "HIV Infection and Sexually Transmitted Diseases." In D. T. Jamison, W. H. Mosley, A. R. Measham, and J. L. Bobadilla, eds. *Disease Control Priorities in Developing Countries*. New York: Oxford University Press.
- Pauly, M. V. 1995. "Valuing Health Care Benefits in Money Terms." In F.A. Sloan, ed., *Valuing Health Care*. Cambridge: Cambridge University Press.
- Petry, K. U., and H. Kingu. 1996. "HIV Infection among Pregnant Women in Lindi, Tanzania, 1989–1993." *International Journal of STD & AIDS* 7:265–68.
- Philipson, T. J. 2000. "Economic Epidemiology and Infectious Diseases." In A. J. Cuyler and J. P. Newhouse, eds., *Handbook of Health Economics*. Amsterdam: North Holland.
- Philipson, T. J., and R. A. Posner. 1993. *Private Choices and Public Health: The AIDS Epidemic in an Economic Perspective*. Cambridge, Mass.: Harvard University Press.
- Prescott, N. M. 1987. "The Economics of Schistosomiasis Chemotherapy." *Parasitology Today* 3:21–24.
- Roth, D. L., K. E. Stewart, O. J. Clay, A. van der Straten, E. Karita, and S. Allen. 2001. "Sexual Practices of HIV Discordant and Concordant Couples in Rwanda: Effects of a Testing and Counselling Programme for Men." *International Journal of STD & AIDS* 12:181–88.
- Ryder, R. W., V. L. Batter, M. Nsuami, N. Badi, L. Mundeke, B. Matela, M. Utshudi, and W. L. Heyward. 1991. "Fertility Rates in 238 HIV-1-Seropositive Women in Zaire Followed for 3 Years Post-Partum." *AIDS* 5:1521–27.
- Salant, S. W. 2003. "Same Infection. Same Time. Same Antibiotic?" In R. Laxminarayan, ed., *Battling Resistance to Antibiotics and Pesticides*. Washington, D.C.: Resources for the Future.
- Serwadda, D., R. H. Gray, M. J. Wawer, R. Y. Stallings, N. K. Sewankambo, J. K. Konde-Lule, B. Lainjo, and R. Kelly. 1995. "The Social Dynamics of HIV Transmission as Reflected through Discordant Couples in Rural Uganda." *AIDS* 9:745–50.
- Sethi, S. P., and P. W. Staats. 1978. "Optimal Control of Some Simple Deterministic Epidemic Models." *Journal of the Operational Research Society* 29:129–36.
- Stiglitz, J. E. 1997. "Introduction." In World Bank, *Contributions to AIDS*. New York: Oxford University Press.

- . 2000. *Economics of the Public Sector*. New York: Norton.
- Temmerman, M., J. Ndinya-Achola, J. Ambani, and P. Piot. 1995. "The Right Not to Know HIV-Test Results." *Lancet* 395:969–70.
- Van der Straten, A., R. King, O. Grinstead, A. Serufilira, and S. Allen. 1995. "Couple Communication, Sexual Coercion and HIV Risk Reduction in Kigali, Rwanda." *AIDS* 9:935–44.
- Viscusi, W. K. 1992. *Fatal Tradeoffs: Public and Private Responsibilities for Risk*. New York: Oxford University Press.
- United Nations and World Health Organization. 1991. *The AIDS Epidemic and its Demographic Consequences*. New York and Geneva: United Nations and World Health Organization.
- Whitaker, L., and A. M. Rentin. 1992. "A Theoretical Problem of Interpreting the Recently Reported Increase of Homosexual Gonorrhoea." *European Journal of Epidemiology* 8:187–91.
- Wickwire, K. 1977. "Mathematical Models for the Control of Pests and Infectious Diseases: A Survey." *Theoretical Population Biology* 11:182–238.
- Wiemer, C. 1987. "Optimal Disease Control through Combined Use of Preventive and Curative Measures." *Journal of Development Economics* 25:301–19.

Vertical line on the left side of the page.

Globalization and the Labor Market

Martín Rama

Does globalization affect labor market outcomes? Can labor market policies mitigate or offset the effects? Would these policies have important side effects on efficiency? This article addresses these questions through an analytical survey of the literature, including several studies under preparation. Some of the studies use new cross-country databases of wages and other labor market indicators. Although all the answers should be considered tentative, some patterns emerge. Different aspects of globalization have different consequences. In the short run wages fall with openness to trade and rise with foreign direct investment. But after a few years the effect of trade on wages becomes positive. Foreign direct investment also increases (substantially) the returns to education. Social protection programs are effective in reducing inequality. Minimum wages, public sector employment, and core labor standards are not. Between these two extremes, collective bargaining works mainly for the middle class. Social protection programs do not adversely affect efficiency, but high public sector employment and trade union membership are associated with weaker performance in the context of adjustment.

Integration with world markets bears the promise of prosperity for developing and transition economies, but it could also be a source of increased hardship. In principle, the unleashing of market forces associated with globalization should increase productivity and possibly economic growth. But many fear that the pendulum will swing too far toward efficiency, to the detriment of equity. Low wages and limited workers' rights could be necessary to attract foreign investment and increase export market shares, and these would be to the disadvantage of workers as a group. Moreover, inequality among workers could increase if greater economic integration benefits those with the skills needed to adjust to the new technologies and organizational structures, while the rest are left behind. These are the fears that lie behind such popular concepts as the "race to the bottom" and the "digital divide."

This pessimistic view contrasts sharply with predictions by economists. One of the most widely used international trade models suggests that lower tariffs and transport

costs should push each country to specialize in the production of the goods for which it has a comparative advantage. In relative terms, skilled labor is the abundant factor in the industrial world and unskilled labor the abundant factor in the developing world. Globalization should therefore be associated with an increase in the relative demand for skilled labor in industrial countries and unskilled labor in developing economies. Based on this view, economic integration could increase inequality within industrial countries—or increase unemployment if labor regulations prevented a downward adjustment in the wages of unskilled workers (Wood 1994). But it should reduce inequality within developing economies. In fact, the labor-intensive growth associated with greater openness and deregulation is often seen as one of the main avenues toward poverty reduction (World Bank 1990).

Both views about the effects of globalization are plausible, and both can be articulated under the form of rigorous economic models. That is why their relevance can be assessed only on empirical grounds. Unfortunately, the available evidence is scattered. The experience of East Asia in recent decades suggests that export-oriented growth in labor-intensive economies is indeed associated with a dramatic reduction in poverty (World Bank 1993). But there has been no systematic decline in inequality in East Asia, as the simple international trade model would have predicted. In fact, inequality has increased dramatically in China. Several studies suggest that wage inequality increased in Latin American countries as they liberalized foreign trade. Although the interpretation of the findings of these studies is open to debate, all the studies show an increase in the returns to skill over periods of economic reform and increased openness, despite differences in data sets and methodologies.¹ This ambiguity in the empirical evidence warrants a closer look at the merits of the two views already outlined.

This article analyzes in more depth the relationship between globalization and the labor market. It addresses three questions: Does globalization have systematic effects on labor earnings? Are domestic labor market policies an effective tool for mitigating or offsetting these effects? Would labor market policies have important side effects on economic efficiency? These questions are addressed mainly from the perspective of developing and transition economies. In these countries wage earners do not necessarily represent a majority of the labor force, as many workers are self-employed or engaged in household enterprises. In addition, the ability to enforce labor market policies is limited, and efforts to enforce such policies often do not reach wage earners in the informal sector.

Although the article presents some new empirical evidence, it should be seen as an analytical survey. The new evidence stems from a series of studies in preparation or recently completed (a more detailed discussion of the data sets and methodologies used in each case can be found in the studies). By putting together some of the new evidence as well as the main findings from other studies, the article tries to provide a consistent story line. Although the answers are tentative, a story line emerges.

Does Globalization Affect Labor Market Outcomes?

Claims that globalization leads to a “race to the bottom”—or, conversely, to greater demand for unskilled labor—implicitly assume that the equilibrium level of labor earnings is affected by international trade and foreign direct investment. But for the large fraction of the labor force that is self-employed or works in household enterprises in many developing economies, globalization is unlikely to have a direct, visible impact on earnings. International trade and foreign direct investment usually involve enterprises that employ salaried workers and often operate in the formal sector. Access to world markets, import penetration, or capital inflows should therefore have the most direct impact on the demand for labor by those enterprises. Competition in the labor market could subsequently translate this direct impact into changes in expenditure or earnings among the self-employed or unpaid family workers. But the first-order effect of globalization should be noticeable in the level of wages, particularly formal sector wages.

A recently released data set, assembled by Freeman and Oostendorp (2000) based on the October Inquiry, can be used to analyze the links between globalization and wages. The October Inquiry, conducted by the International Labour Office (ILO) since 1924, requests data on wages by occupation from governments around the world. These data have rarely been used for research because they suffer from comparability problems. Occupations (teacher, bricklayer, and the like) are defined quite precisely. But depending on the country and occupation, the figures submitted by governments might refer to wages or to earnings. They might be legal minimums or maximums or actual averages or medians. Figures could refer to men, women, or both. The reference period might be the hour, day, week, month, or something else.

Freeman and Oostendorp recalibrated these data to make them comparable, assuming a stable relationship between different wage measures. The data used in this article correspond to the average monthly wages of men, measured in current dollars adjusted for purchasing power parity (PPP). The earlier observations correspond to 1983, but the occupations for which data are available vary across countries and years. An effort is under way to extend the recalibration back to 1953.

To assess whether globalization is associated with a “race to the bottom,” wage levels can be regressed on openness indicators.² In table 1 the explained variable is the log of wages by occupation, based on the data set assembled by Freeman and Oostendorp. In all the regressions reported in this table, there is a maximum of one observation per country, year, and occupation. A key explanatory variable is the ratio of trade to gross domestic product (GDP), but three other openness indicators are considered. One is simply a variant of this ratio, in which PPP dollars (rather than current dollars) are used to measure the numerator and the denominator. Another indicator, constructed by Sachs and Warner (1995), focuses on economic policies. An economy is considered open if it meets five criteria: low tariffs, limited coverage of

Table 1. Wages by Occupation and Openness

| Explanatory variable | Explained variable: log of wages by occupation (PPP dollars) | | | | | | | |
|---|--|-----------------------|-------------------|-----------------------|-----------------------|----------------------|--------------------|----------------------|
| | | | | | | | | |
| Ratio of trade to GDP (based on current dollars) | -0.125*** (-7.41) | -0.284*** (-14.74) | | | | | | |
| Ratio of trade to GDP (based on PPP dollars) | | | -0.017 (-0.78) | -0.318*** (-11.73) | | | | |
| Open economy (yes = 1) | | | | | -0.131*** (-14.18) | -0.046*** (-4.25) | | |
| Ratio of foreign direct investment to GDP (based on current dollars) | | | | | | | 1.062*** (8.71) | 1.178*** (8.38) |
| Log of GDP per capita (PPP dollars) | | 0.662*** (37.92) | | 0.794*** (25.80) | | 0.718*** (32.61) | | 0.691*** (39.26) |
| Political liberty index (scale: 0-1) | | 0.103*** (6.53) | | 0.110*** (4.82) | | 0.065*** (3.63) | | 0.102*** (6.36) |
| Economic liberty index (scale: 0-1) | | -0.129*** (-3.39) | | -0.052 (-0.97) | | -0.193*** (-4.10) | | -0.212*** (-5.45) |
| Country effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Occupation effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| R ² | 0.438 | 0.456 | 0.465 | 0.465 | 0.446 | 0.450 | 0.437 | 0.454 |
| F-test | 261.1 | 227.9 | 196.3 | 166.7 | 212.0 | 192.3 | 246.7 | 219.2 |
| Number of observations | 59,292 | 48,934 | 38,852 | 33,647 | 45,884 | 41,557 | 56,317 | 47,506 |

Note: All regressions are estimated using fixed effects. All explanatory variables are lagged one year, with *t*-values reported in parentheses. Three asterisks denote statistical significance at the 1 percent level.

Source: Author's calculations based on work in progress by Freeman and others (2001). Data on wages by occupation are from Freeman and Oostendorp (2000); data on the ratio of trade to GDP in PPP terms are from the Penn World Tables (Heston and others 2002); data on openness are from Sachs and Warner (1995); data on political liberties are based on the Polity III index as rescaled by Rodrik (1999); the economic liberty index is produced by the Fraser Institute (Easton and Walker 1992), rescaled; all other data are from the World Bank.

nontariff barriers, no marketing boards, no central planning, and a small or nonexistent black market premium for the exchange rate. The last openness indicator considered is the ratio of foreign direct investment to GDP; for this indicator both the numerator and the denominator are measured in current dollars.

All the regressions reported in table 1 are estimated using fixed effects. This method is the same as including dummy variables for each country, year, and occupation. The coefficients on these dummy variables remove the effect of unobservable characteristics, so that the estimated coefficient on an openness indicator measures the impact of the change in exposure to world markets on the change in wage levels. The implicit hypothesis when assessing the sign and significance of this coefficient is that globalization has a systematic effect, in one direction or the other, on all the wages considered in the analysis. A significantly negative coefficient should be interpreted as evidence that wages are lower in more open economies, everything else equal—or, equivalently, that greater openness will exert a downward pressure on wages.

Regressions of this sort raise several important problems. Among the most important are causality, omitted variables, and structural stability. If a relationship is found, it could be difficult to tell whether openness has an effect on wages, or the other way around. To address this problem, all the explanatory variables in table 1 were lagged one year, meaning that the key coefficient measures the relationship between openness in any one year and the wage level in the year after.

Because of omitted variables, it could well be the case that some other economic force explains both openness and the wage level, so that the estimated relationship is misleading (and probably biased). To address this problem, all regressions were rerun with the addition of three control variables, accounting for the level of economic development, political liberty, and economic liberty. The inclusion of political liberty was based on previous evidence on its relevance when explaining wage levels across countries (Rodrik 1999). Economic development is measured by the log of GDP per capita in PPP dollars, political liberty by the Polity III index, and economic liberty by the Fraser index. Other regressions (not reported in this article) consider alternative development indicators, such as the educational attainment of the population or the urbanization rate. Lags of one year were used for these variables as well. The estimated impact of globalization on wages did not change substantially when these additional controls were included in the regressions, but its statistical significance fell as the number of usable observations declined.

There would be structural stability problems if the relationship between globalization and wages differed across subsets of countries. In particular, the simple international trade model sketched in the introduction predicts that the impact of openness on wages should differ between industrial and developing economies because of the difference in skill composition between their labor forces. All regressions were therefore rerun with industrial countries excluded from the sample. The estimated coefficients

remained very similar in all cases (the results are not reported herein). But their statistical significance declined because of the smaller sample.

The results reported in table 1 suggest that in the short run globalization has a mixed impact on wages. Openness to trade, as captured by trade flows or trade liberalization policies, is associated with lower wages by occupation. Except in one specification, all the effects are statistically significant. In addition, the size of the effects appears to be considerable. In the specifications that control for the level of economic development and political and economic liberty, a 20-percentage-point increase in the ratio of trade to GDP leads to a 5–6 percent decline in wages. According to the indicator constructed by Sachs and Warner, the impact is similar when an economy opens up. In contrast, foreign direct investment appears to have a positive impact on wages. When foreign direct investment as a share of GDP increases by one percentage point, wages grow by roughly 1 percent.

A simple way to test whether globalization also leads to increased inequality among workers is to repeat the analysis in table 1 using indicators of wage dispersion as the explained variable. In table 2 the dispersion indicator is the standard deviation of the log of wages by occupation based on the data set assembled by Freeman and Oostendorp. This standard deviation can be interpreted as the typical gap, in relative terms, between wages in any occupation and the average wage. The higher the standard deviation, the greater the inequality among workers. The regressions in table 2 include a maximum of one observation per country and year.

Returns to education provide an alternative and potentially more telling indicator of wage dispersion. The larger the returns, the wider the earnings gap between skilled and unskilled workers. A variety of studies around the world have tried to estimate the impact of an additional year of education on labor earnings, using individual records from household or labor force surveys. Many of these studies rely on a Mincerian earnings function, explaining the log of earnings as a function of the number of years of schooling, the number of years of potential work experience, and the square of the years of potential work experience. These studies use data sources of different coverage and quality, with inconsistent measures of labor earnings. The Mincerian equations they report include different control variables and involve different subsets of workers (for example, private sector workers, urban workers, or male workers). But as long as the “noise” in the estimated returns to education is not correlated with openness or other explanatory variables in the regressions, those estimated returns provide a valuable source of information for studying the impact of globalization on labor market outcomes.

In table 3 the dispersion indicator is the percentage increase in labor earnings associated with one additional year of education. This indicator is based on a new (and preliminary) data set on returns to education that is being collected as part of work in progress, building on previous efforts by Psacharopoulos (1994). The data set reports the estimated coefficient for the years of schooling in the “preferred”

Table 2. Dispersion of Wages by Occupation and Openness

| Explanatory variable | Explained variable: standard deviation of the log of wages by occupation | | | | | | | |
|---|--|---------|---------|---------|---------|---------|--------|---------|
| | | | | | | | | |
| Ratio of trade to GDP (based on current dollars) | 0.052* | 0.025 | | | | | | |
| | (1.88) | (0.68) | | | | | | |
| Ratio of trade to GDP (based on PPP dollars) | | | -0.028 | -0.056 | | | | |
| | | | (-0.76) | (-1.04) | | | | |
| Open economy (yes = 1) | | | | | -0.008 | -0.006 | | |
| | | | | | (-0.46) | (-0.29) | | |
| Ratio of foreign direct investment to GDP (based on current dollars) | | | | | | | 0.295 | -0.171 |
| | | | | | | | (1.31) | (-0.60) |
| Log of GDP per capita (PPP dollars) | | -0.052 | | 0.012 | | -0.045 | | -0.055 |
| | | (-1.46) | | (0.20) | | (-0.99) | | (-1.50) |
| Political liberty index (scale: 0–1) | | 0.034 | | 0.015 | | 0.021 | | 0.032 |
| | | (1.10) | | (0.35) | | (0.60) | | (1.02) |
| Economic liberty index (scale: 0–1) | | -0.067 | | -0.176* | | -0.174* | | -0.063 |
| | | (-0.91) | | (-1.70) | | (-1.84) | | (-0.81) |
| Country effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of occupations considered | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| R ² | 0.001 | 0.303 | 0.000 | 0.001 | 0.012 | 0.254 | 0.019 | 0.291 |
| F-test | 1.95 | 1.35 | 1.37 | 1.30 | 1.40 | 1.57 | 2.53 | 1.31 |
| Number of observations | 858 | 642 | 575 | 445 | 661 | 535 | 784 | 611 |

Note: All regressions are estimated using fixed effects. All explanatory variables are lagged one year, with *t*-values reported in parentheses. An asterisk denotes statistical significance at the 10 percent level.

Source: Author's calculations based on work in progress by Freeman and others (2001). Data on wages by occupation are from Freeman and Oostendorp (2000); data on the ratio of trade to GDP in PPP terms are from the Penn World Tables (Heston and others 2002); data on openness are from Sachs and Warner (1995); data on political liberties are based on the Polity III index as rescaled by Rodrik (1999); the economic liberty index is produced by the Fraser Institute (Easton and Walker 1992), rescaled; all other data are from the World Bank.

Table 3. Returns to Education and Openness

| Explanatory variable | Explained variable: additional earnings per year of schooling (percent) | | | | | | | |
|---|---|---------|--------|---------|--------|---------|---------|---------|
| | | | | | | | | |
| Ratio of trade to GDP (based on current dollars) | 3.151* | 2.420 | | | | | | |
| | (1.88) | (0.90) | | | | | | |
| Ratio of trade to GDP (based on PPP dollars) | | | 2.545 | 0.475 | | | | |
| | | | (1.45) | (0.16) | | | | |
| Open economy (yes = 1) | | | | | 0.258 | -0.088 | | |
| | | | | | (0.41) | (-0.10) | | |
| Ratio of foreign direct investment to GDP (based on current dollars) | | | | | | | 50.91** | 54.56** |
| | | | | | | | (2.45) | (2.03) |
| Log of GDP per capita (PPP dollars) | | -3.091* | | -0.412 | | -3.229* | | -3.085 |
| | | (-1.68) | | (-0.19) | | (-1.69) | | (-1.62) |
| Political liberty index (scale: 0-1) | | -2.224 | | -2.382 | | -3.260* | | -2.289 |
| | | (-1.43) | | (-1.52) | | (-1.93) | | (-1.46) |
| Economic liberty index (scale: 0-1) | | -0.628 | | -1.336 | | -1.003 | | -1.094 |
| | | (-0.19) | | (-0.34) | | (-0.27) | | (-0.33) |
| Country effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| R ² | 0.008 | 0.072 | 0.011 | 0.087 | 0.028 | 0.083 | 0.026 | 0.110 |
| F-test | 1.94 | 1.80 | 1.70 | 2.10 | 1.78 | 2.04 | 2.24 | 2.07 |
| Number of observations | 507 | 352 | 462 | 318 | 503 | 340 | 432 | 344 |

Note: All regressions are estimated using fixed effects. All explanatory variables are lagged one year, with *t*-values reported in parentheses. Significant coefficients at the 10 and 5 percent levels are indicated by one and two asterisks, respectively.

Source: Author's calculations based on work in progress by Freeman and others (2001). Data on the ratio of trade to GDP in PPP terms are from the Penn World Tables (Heston and others 2002); data on openness are from Sachs and Warner (1995); data on political liberties are based on the Polity III index as rescaled by Rodrik (1999); the economic liberty index is produced by the Fraser Institute (Easton and Walker 1992), rescaled; all other data are from the World Bank.

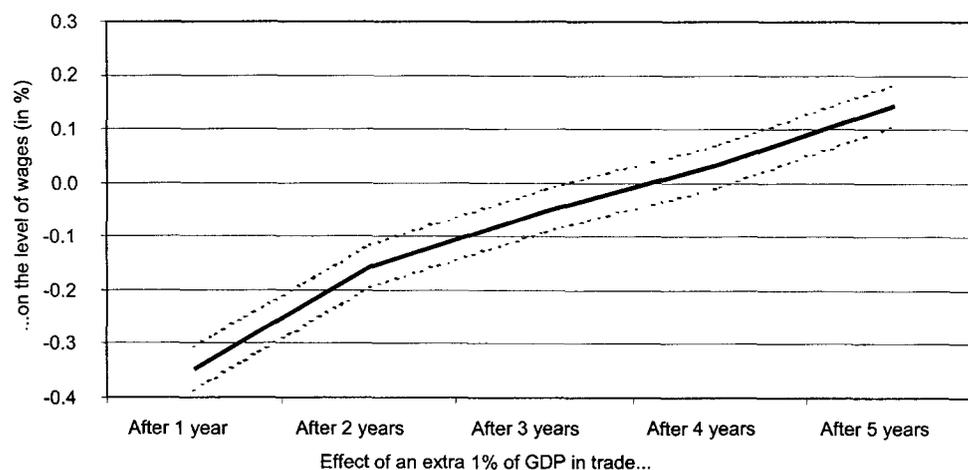
specification in each of the studies available. For some countries and years no estimate is available, perhaps because no household or labor force survey was carried out in that year, because no study was conducted, or simply because the Mincerian approach was not used. For some other countries there may be many estimates for a single year. As long as these estimates come from different studies, they are all included in the data set. As a result, there can be several observations per country and year in the regressions in table 3, but missing values are common.

The specification of the regressions in tables 2 and 3 is the same as that in table 1, with one exception. In the data set assembled by Freeman and Oostendorp the number of occupations for which wage data are available varies substantially across countries and years. The standard deviation of the log of wages is sensitive to the number of occupations for which wage data are available. At the limit, if there was only one observation for a specific country and year, the estimated standard deviation would be equal to zero because there would be no difference between the only observed wage and the average wage. To correct the potential effects of this measurement error, the regressions in table 2 include the number of occupations considered in each country and year as an additional explanatory variable.

According to the results reported in table 2, globalization has no impact on the dispersion of wages across occupations. The sign of the coefficient of interest varies across specifications and is almost never statistically significant. There also appears to be no relation between openness to trade and returns to education. But according to the results reported in table 3, foreign direct investment is associated with a substantial increase in the labor market premium for skill. If interpreted literally, the coefficients in table 3 mean that a one-percentage-point increase in foreign direct investment as a share of GDP is associated with a five-percentage-point increase in the returns to an additional year of schooling. This suggests that the "digital divide" effect, if it does exist, operates through the technology embedded in foreign capital rather than directly through foreign trade. This finding is consistent with microeconomic evidence on the effects of *maquiladoras* in Mexico (Feenstra and Hanson 1997).

The results in tables 1, 2, and 3 refer to impacts over one year. Over longer periods the effects of globalization could be quite different. Figures 1 and 2 summarize the results of a series of regressions that are in the spirit of those in table 1 except that they involve longer time lags between the openness indicators and the level of wages. These regressions include both the ratio of trade to GDP (measured in dollars) and the ratio of foreign direct investment to GDP (also measured in dollars) among their explanatory variables. They also include controls for the level of development, political liberty, and economic liberty. Figure 1 reports the coefficients associated with the ratio of trade to GDP for time lags ranging from one year (as in table 1) to five years. Figure 2 reports the coefficients for the ratio of foreign direct investment to GDP over the same time range. The solid lines in the two figures correspond to the point estimate; the dotted lines indicate the 95 percent confidence interval. The point

Figure 1. Effect on the Wage Level of a One-Percentage-Point Increase in Trade as a Share of GDP



Note: This figure reports the results of a regression explaining the log of wages (in PPP dollars) as a function of the ratio of trade to GDP (in dollars). There is one observation per country, occupation, and year over the period 1983–98. The regression includes fixed effects for all countries, occupations, and years. It also controls for the ratio of foreign direct investment to GDP. The solid line reflects the point estimate of the coefficient of interest. The dotted lines indicate the 95 percent confidence interval.

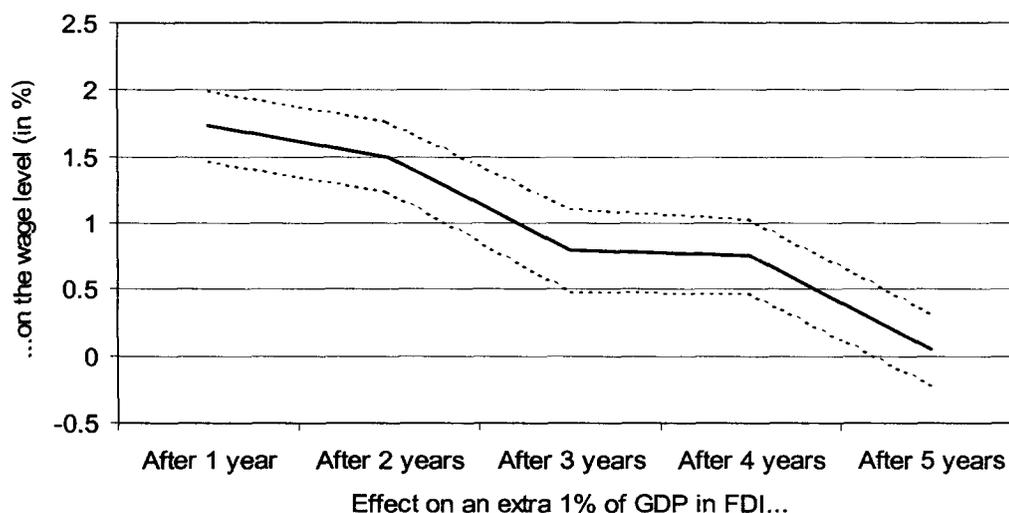
Source: Author's calculations based on work in progress by Freeman and others (2001). Data on wages are from Freeman and Oostendorp (2000), and data on other variables from the World Bank.

estimates with a one-year lag differ slightly from the coefficients reported in table 1 because the two openness indicators were entered among the explanatory variables simultaneously in the regressions underlying figures 1 and 2.

Figure 1 shows that trade openness has a significantly negative impact on wages at first, but the effect evolves steadily and becomes positive in the medium term. The change in sign occurs after three years when the openness indicator is the ratio of trade to GDP. It occurs after five years if the indicator is the composite policy variable constructed by Sachs and Warner (1995). Figure 2 shows that the impact of foreign direct investment on wages remains positive but becomes statistically insignificant after five years.

In the longer term, openness should have an impact on wages through its overall effect on economic development. It is generally accepted that openness to trade is associated with higher growth in output per capita (see, for example, Sala-i-Martin 1997). The level of wages is, in turn, highly responsive to economic development. In all the regressions that control for economic development in table 1, the coefficient multiplying the log of income per capita is positive and highly significant. If openness is good for output growth, it should also be good for wage growth over the long run.

Figure 2. Effect on the Wage Level of a One-Percentage-Point Increase in Foreign Direct Investment as a Share of GDP



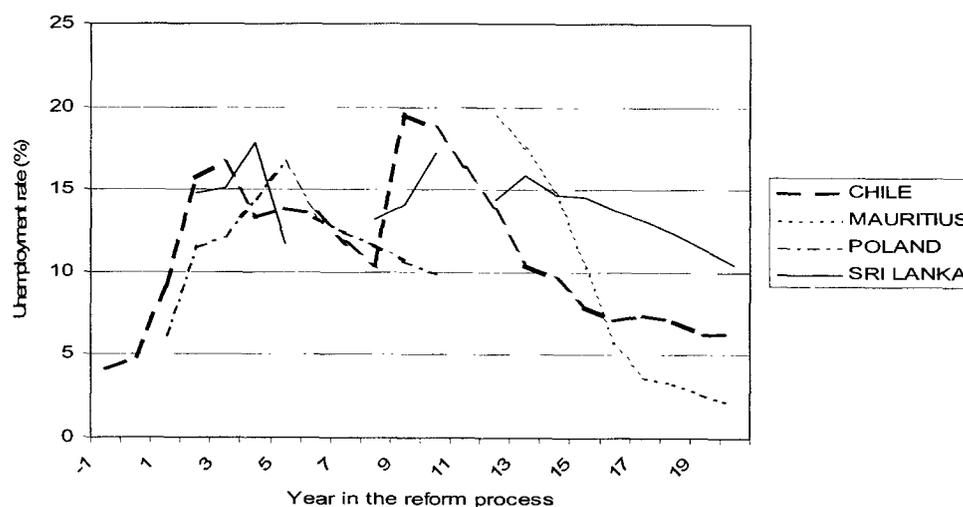
Note: This figure reports the results of a regression explaining the log of wages (in PPP dollars) as a function of the ratio of foreign direct investment to GDP (in dollars). There is one wage observation per country, occupation, and year over the period 1983–98. The regression includes fixed effects for all countries, occupations, and years. It also controls for the ratio of trade to GDP. The solid line reflects the point estimate of the coefficient of interest. The dotted lines indicate the 95 percent confidence interval.

Source: Author's calculations based on work in progress by Freeman and others (2001). Data on wages are from Freeman and Oostendorp (2000), and data on other variables from the World Bank.

For some workers, however, trade liberalization could lead to much hardship in the short run, even more than that suggested by the negative impact of openness on wages reported in table 1. Many jobs in protected (presumably inefficient) industries could be lost as a result of trade liberalization and market deregulation. For workers in those industries the resulting decline in earnings would exceed the decline in wages that can be predicted based on the estimated regressions. Optimistically, one could argue that laid-off workers will find other jobs. But jobs in protected industries are usually “better” jobs, probably because of some form of rent sharing (Azam and Ris 1995; Morisson 1994). Moving to another job would presumably entail the loss of the “protection” premium.

Moreover, the experience of some “successful” reformers shows that unemployment rates may remain stubbornly high for long periods. In Chile, Mauritius, Poland, and Sri Lanka after a major liberalization effort, unemployment rates climbed to double digits and stayed there for more than a decade (figure 3). Of course, this outcome can always be blamed on adverse shocks and economic policy

Figure 3. Unemployment Rate in Successful Reformers



Note: For the purposes of the figure, economic reforms start with the end of socialism in Chile (1973) and Poland (1991), with the adoption of the export processing zone regime in Mauritius (1970), and with trade liberalization in Sri Lanka (1977).

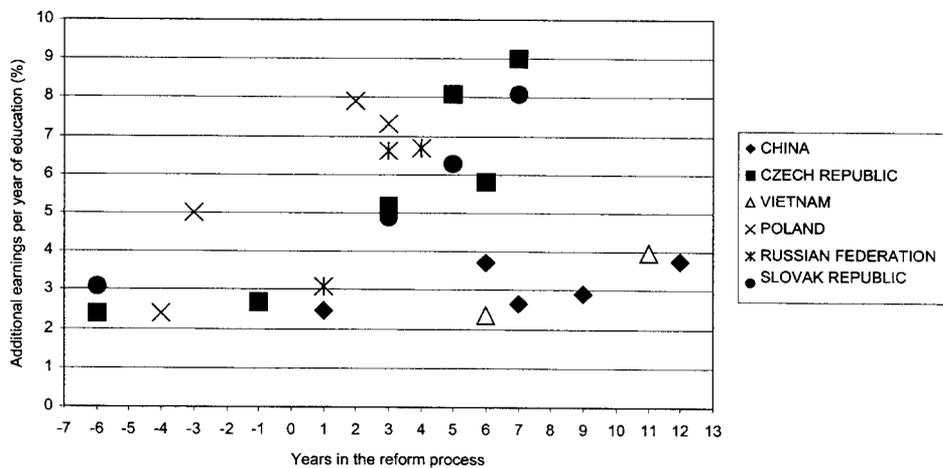
Source: Reproduced from World Bank (2001).

mistakes. But this criticism has limits, because these four countries are success stories in their own regions. Other developing economies would probably find it difficult to do much better.

Similarly, there are countries in which economic reforms (including greater openness to trade) could be associated with a sharper increase in returns to education than that suggested by the econometric results presented here. The experience of formerly socialist economies is revealing. These economies were characterized by highly compressed salary structures under central planning. But market forces widened the wage distribution, introducing a substantial premium for skill. This trend is illustrated in figure 4, which shows returns to education in formerly socialist countries in the years following the start of the transition to a market economy. For some of these countries the observed increase in returns to education amounts to a widening of the wage gap by 80 percentage points between workers with a college education and those with a primary education. Freeman and Oostendorp (2000) also report a marked increase in the dispersion of wages across occupations in formerly socialist economies.

In sum, openness should benefit workers in the long run and has the potential to benefit skilled workers even in the short run. But it may also penalize wage earners who are unskilled or work in protected activities. The net impact depends on

Figure 4. Returns to Education in Selected Transition Economies



Note: The figures reported are the coefficients multiplying the years of education in Mincerian earnings functions estimated using microeconomic data.

Source: Reproduced from World Bank (2001).

whether globalization is associated only with more trade or also with more foreign direct investment. Greater foreign direct investment has the potential to raise wages rapidly, but it may also widen the wage gap between skilled and unskilled workers. Another way to summarize the findings of this section is to say that openness may create winners and losers. Skilled workers are the most likely winners. Young workers, who can expect to benefit from wage growth in the long run, should also be among the winners. Unskilled and old workers, especially those in protected industries, could be the losers. This heterogeneity of outcomes has important political economy implications, analyzed in the section on the efficiency effect of labor market policies.

Can Labor Market Policies Mitigate the Effects of Globalization?

Because globalization has effects on labor market outcomes, it is only natural to consider policies specifically related to the labor market as a mechanism for mitigating those effects and, more broadly, their potential impact on inequality. In recent years there has been much debate about the merits and demerits of minimum wages, mandated benefits, collective bargaining, job security, and public sector employment in developing economies. There are even proposals to make the link between

labor market policies and openness explicit by including compliance with labor standards among the rules of the multilateral trading system. Those standards usually take the form of the seven core conventions of the ILO. A pertinent question is whether labor market policies, including core labor standards, are relevant in countries where only a fraction of the labor force is salaried and the government has limited ability to enforce regulations.

A first step in addressing this question is to review the available evidence on the effects of labor market policies and institutions on economic performance. Although the literature was limited to industrial countries until a short time ago, there are now many labor market studies on developing economies. Many of these studies are based on microeconomic data, which makes them more credible. The literature can be used to assess the potential of different labor market interventions to counter the effects of globalization.

Studies on minimum wages suggest that their impact is likely to be small. In many developing economies it is difficult to find a sizable spike in the distribution of wage earnings for full-time workers around the minimum wage (Maloney and Nuñez 2003). Even in countries with a relatively high administrative capacity, the fraction of full-time workers earning less than the minimum wage is roughly the same in the formal and informal sectors (Gindling and Terrell 1995). Not surprisingly, employment effects of minimum wages are difficult to find except when the minimum wage is very high compared with the average wage, as in Colombia (Bell 1997). An interesting “natural” experiment was the doubling of minimum wages in real terms in Indonesia in the early 1990s. During this period the elasticity of average wages to minimum wages at the provincial level was roughly 10 percent. Though there was a mild decline in total wage employment, the disemployment effect was substantial in small establishments (Rama 2001).

Studies on mandated benefits indicate that their cost might be shifted to workers, without modifying total compensation much. That workers “pay” for their benefits through lower cash earnings is not surprising when those benefits are perceived as a form of compensation. More strikingly, there also appears to be a one-to-one shift for benefits whose value is less clear from the workers’ perspective. That was what happened with the reform of the social security system in Chile, when a reduction in contributions was offset by an increase in wage earnings even though benefits in the old system were only loosely related to contributions (Gruber 1997). Conversely, the cost of some benefits may be shifted to employers even if they are paid in cash and are therefore fully fungible with wages. That is what happened in Ecuador in the public sector and in unionized activities (MacIsaac and Rama 1997).

Trade unions seem to be effective in raising the wages of their members, but the union wage premiums are probably smaller in developing than in industrial countries. The reason may be that the labor movement is “subordinate” in many developing economies, where some trade unions were even instrumental in enacting wage

freezes during periods of economic adjustment (Nelson 1991). Thus the estimated union wage premiums range from negligible in Senegal to small in Mexico (Terrell and Svejnar 1989; Panagides and Patrinos 1994). It has been argued that South Africa is an exception, with an extraordinarily high union wage premium (Mwabu and Schultz 2000). But this high premium probably reflects self-selection more than bargaining power. An in-depth analysis of the South African data suggests that the wage premium is roughly 15 percent, much the same as in the United States. This analysis also shows that trade unions extend part of this premium to nonunionized workers in activities covered by industrial councils (Butcher and Rouse 2001).

The impact of separation costs is more controversial. A few countries have mandated job security, meaning that some employers (typically large ones) cannot dismiss a worker without explicit approval from the government. Because this approval is political, employers may be reluctant to hire workers under permanent contracts, for fear of being unable to fire them in bad times. There is evidence that the tightening of job security regulations depressed labor demand in India and Zimbabwe (Fallon and Lucas 1991). In most developing economies, however, separation costs take the form of mandatory severance pay: the employer must pay an indemnity to workers who lose their job through no fault of their own. This indemnity usually depends on wage and seniority. Based on cross-country data, it has been claimed that separation costs of this kind reduce employment (Heckman and Pagés 2000). But microeconomic data from Peru suggest that workers may “pay” for some of their separation package through lower wages (MacIsaac and Rama 2001).

Relatively few studies have looked at the impact of public sector employment on labor market outcomes. Where the formal sector is very small, as in Sub-Saharan Africa, public sector wages could have a substantial impact on private sector wages (Rama 2000). The fact that public sector jobs are “better” than their private sector counterparts could be at the root of “queuing unemployment,” as in the Arab Republic of Egypt and Sri Lanka (Assaad 1997; Rama 2003). Public sector jobs could also serve as a form of insurance, especially in countries where intrahousehold transfers are substantial. This possibility has been suggested by Rodrik (2000), based on cross-country data.

This admittedly brief review of the literature suggests that labor market policies could have an impact on labor market outcomes and, more broadly, on inequality. But this impact is probably small and its direction is unclear. Higher minimum wages and stronger trade unions could raise the earnings of those who manage to keep their jobs but reduce labor demand and depress the earnings of those in the informal sector. Formal sector workers usually belong to the “middle class” in developing economies, whereas informal sector workers are more likely to be poor. More generous mandated benefits and higher separation costs could have a similar impact as long as workers do not fully pay for their coverage. Public sector employment

could benefit the middle class rather than the poor. Still, all these effects might be too small to make a difference.

Cross-country analysis can shed light on the impact of labor market policies on inequality more broadly. Several studies have used the database assembled by Deininger and Squire (1996) to explore the links between conventional indicators of inequality, such as the Gini coefficient, and a set of country characteristics. Among the explanatory variables used in those studies are the average educational attainment of the population and indicators of civil liberties, financial depth, and the concentration of land ownership (Li and others 1998). Labor market indicators have been largely ignored. Bourguignon and Morrisson (1998) appear to be the only researchers to have included a labor market variable among the determinants of inequality. But that variable—the ratio of labor productivity in manufacturing to that in agriculture—measures a labor market outcome rather than a labor market policy.

To redress this omission, the inequality indicators in the database assembled by Deininger and Squire can be paired with labor market indicators measuring the extent of government intervention in the labor market. Six inequality indicators are considered in what follows. Five refer to the share of each population quintile in total consumption or earnings, measured in logs. When data on both consumption and earnings are available, the consumption data are preferred. The sixth indicator is the Gini index for the distribution of consumption or earnings, also in logs. The actual data are from a revised and updated version of the Deininger and Squire data set.³

The labor market indicators considered are from a broader data set being constructed by Rama and Artecona (2002) and are computed as averages over five-year periods. They include the number of ILO conventions ratified by the country (all conventions and core conventions). They also include the most general minimum wage, measured as a percentage of the average labor cost per worker in manufacturing (the average labor cost is from establishment surveys and censuses). The extent of social protection is measured by the social security contributions paid by employers and employees as a percentage of the gross wage, by the fraction of GDP channeled through the social security system, and by the mandatory number of days of maternity leave with full pay for a first child born without complications. Three indicators are considered for the importance of collective bargaining: union membership as a percentage of the labor force, the coverage of collective bargaining agreements as a percentage of the salaried workforce, and the number of strikes and lockouts a year. The set of labor market indicators is completed by employment in the general government as a percentage of the labor force and the average government wage as a percentage of the average labor cost per worker in manufacturing.

Crude correlations between inequality indicators and labor market indicators, reported in table 4, suggest that most of the government interventions considered have

Table 4. Correlation Between Inequality and Labor Market Policies

| <i>Labor market indicator</i> | <i>Income or consumption share by quintile</i> | | | | | |
|--|--|---------------|--------------|---------------|----------------|-------------------|
| | <i>Poorest</i> | <i>Second</i> | <i>Third</i> | <i>Fourth</i> | <i>Richest</i> | <i>Gini index</i> |
| ILO conventions ratified | 0.23* | 0.24* | 0.26* | 0.25* | -0.33* | -0.28* |
| Core ILO conventions ratified | 0.03 | 0.03 | 0.04 | 0.05 | -0.11 | -0.09 |
| Minimum wage (% of average) | -0.02 | 0.04 | 0.04 | 0.05 | -0.04 | -0.01 |
| Social security contributions (% of wage) | 0.41* | 0.40* | 0.38* | 0.36* | -0.42* | -0.42* |
| Social security revenue (% of GDP) | 0.51* | 0.52* | 0.53* | 0.52* | -0.53* | -0.50* |
| Days of paid maternity leave | 0.36* | 0.37* | 0.37* | 0.37* | -0.42* | -0.44* |
| Union membership rate (%) | 0.24* | 0.31* | 0.38* | 0.43* | -0.49* | -0.47* |
| Coverage of collective bargaining (%) | 0.52* | 0.55* | 0.56* | 0.58* | -0.57* | -0.58* |
| Number of strikes and lockouts | 0.29* | 0.27* | 0.26* | 0.25* | -0.27* | -0.30* |
| General government (% of labor force) | 0.26* | 0.38* | 0.46* | 0.53* | -0.56* | -0.52* |
| Government wage (% of average) | -0.17 | -0.21 | -0.22 | -0.22 | 0.21 | 0.18 |

Note: The figures in the table are coefficients of correlation between inequality indicators and labor market policies. The number of observations varies from cell to cell depending on data availability. *Significant at the 10 percent level.

Source: Author's calculations based on work in progress by Rama and Ravallion (2001).

considerable redistributive effects. The vast majority of the correlation coefficients in table 4 are statistically significant. Government interventions in the labor market appear to be associated with a lower Gini index and also with a smaller share of consumption or income for the richest population quintile. Labor market interventions appear to increase the share of consumption or income for all other population quintiles, including the poorest one, even though very few among the poor have formal sector jobs. These results are affected by omitted variable bias, however, as the more elaborate analysis in table 5 shows.

The ideal approach to assess the impact of labor market interventions on inequality would be to regress the inequality indicators on the labor market indicators while controlling for other country characteristics. These country characteristics should include determinants of inequality considered in earlier studies, such as the average educational attainment of the population, civil liberties, financial depth, and the concentration of land ownership. A richer analysis should also consider the ratio of trade to GDP, the share of government consumption in GDP, and a variety of interactive and quadratic terms, to capture nonlinear relationships between inequality and its determinants. Unfortunately, labor market indicators are rarely available for the same periods as inequality indicators. Thus this ideal approach would reduce the number of observations that could be used in the regression analysis, sometimes dramatically.

Table 5. Correlation Between “Unexplained” Inequality and Labor Market Policies

| <i>Labor market indicator</i> | <i>Income or consumption share by quintile</i> | | | | | <i>Gini index</i> |
|---|--|---------------|--------------|---------------|----------------|-------------------|
| | <i>Poorest</i> | <i>Second</i> | <i>Third</i> | <i>Fourth</i> | <i>Richest</i> | |
| ILO conventions ratified | 0.30* | 0.26* | 0.24* | 0.23* | -0.25* | -0.26* |
| Core ILO conventions ratified | 0.03 | 0.01 | 0.03 | 0.06 | -0.03 | -0.01 |
| Minimum wage (% of average) | -0.22 | -0.21 | -0.20 | -0.20 | 0.22 | 0.18 |
| Social security contributions (% of wage) | 0.31 | 0.29 | 0.27 | 0.21 | -0.28 | -0.29 |
| Social security revenue (% of GDP) | 0.32* | 0.33* | 0.30* | 0.26* | -0.31* | -0.33* |
| Days of paid maternity leave | 0.22 | 0.18 | 0.12 | 0.09 | -0.06 | -0.05 |
| Union membership rate (%) | 0.14 | 0.14 | 0.16 | 0.18 | -0.17 | -0.10 |
| Coverage of collective bargaining (%) | 0.24 | 0.33 | -0.48 | 0.60* | -0.44 | -0.33 |
| Number of strikes and lockouts | 0.07 | -0.05 | 0.04 | 0.02 | -0.01 | -0.13 |
| General government (% of labor force) | 0.13 | 0.14 | 0.14 | 0.13 | -0.12 | -0.12 |
| Government wage (% of average) | 0.08 | 0.06 | 0.03 | 0.03 | 0.04 | -0.06 |

Note: The figures in the table are coefficients of correlation between “unexplained” inequality indicators and labor market policies. Unexplained inequality is measured as the residual of a regression linking each inequality indicator with a large number of explanatory variables, none of which is directly related to the labor market. The number of observations varies from cell to cell depending on data availability. *Significant at the 10 percent level.

Source: Author’s calculations based on work in progress by Rama and Ravallion (2001).

The alternative considered here is to assess the correlation between labor market indicators and the inequality that is “unexplained” by country characteristics like those just mentioned. This alternative involves two steps. The first is to regress the inequality indicators on all country characteristics, including the interactive and quadratic terms. The second is to compute the correlation between the residuals of those regressions (the unexplained part of inequality) and the labor market indicators. The results of this second step are reported in table 5.

After other country characteristics are controlled for, only two (closely related) labor market indicators have a consistent impact on all inequality indicators: the number of ILO conventions ratified by the country and the share of GDP channeled through the social security system. The regulatory regime promoted by the ILO tends to be protective of workers and to include many benefits that are usually managed by social security administration. In fact, the correlation coefficient between the number of ILO conventions ratified and the share of GDP channeled through social security is quite high. Across all countries and periods covered in the database being constructed by Rama and Artecona (2002), the correlation coefficient reaches almost 0.6.

The only other labor market policy or institution showing a statistically significant correlation with inequality indicators is the share of the labor force covered by collective bargaining agreements. But the effect is significant only for the consumption or income share of the fourth quintile, the quintile to which many formal sector

workers in developing economies are likely to belong. The evidence in the previous section suggests that globalization may adversely affect specific groups of formal sector workers, at least in the short run. But it may affect other population quintiles even more adversely. Focusing on a labor market policy that works only for the second richest group in the population may not be the most effective way to reduce inequality.

All other labor market policies and institutions appear to be ineffective once the other determinants of inequality are controlled for. In particular, minimum wages and core labor standards have no significant impact on any of the inequality indicators. This irrelevance contradicts popular perceptions and casts doubt on the usefulness of including compliance with labor standards among the rules of the multilateral trading system. Public sector employment and public sector wages also appear to be ineffective in reducing inequality. The results in table 5 suggest that only the traditional instruments of social protection work.

Do Labor Market Policies Reduce Efficiency?

Although some labor market policies may be effective in mitigating the adverse impact of globalization on workers, their cost in economic efficiency should not be overlooked. All forms of redistribution entail some deadweight loss, and labor market policies should be no exception. But the deadweight loss may vary substantially across government interventions. To some extent that loss can be inferred from the previous section's review of the microeconomic evidence on the effects of labor market policies in developing economies. This review found an impact on employment for most types of labor market policies. Other things equal, the larger the impact, the bigger the deadweight loss. But the evidence is not precise enough to rigorously compare the efficiency costs of any two government interventions in the labor market. Again, cross-country analysis may be a more promising way to identify the basic regularities.

The long-run impact of labor market policies on economic efficiency can in principle be quantified using Barro regressions. These are relationships linking average growth rates across countries with some of the characteristics of those countries, including their economic policy choices (Barro 1997). This approach has been used to assess the consequences of trade, monetary, financial, fiscal, and other policies. But because of the lack of a reliable cross-country database of labor market indicators, it does not appear to have been applied to labor market policies. Cukierman and others (2001) have tried to fill this gap, using the information compiled by Rama and Artecona (2002). This section reports some of their preliminary estimates.

Besides the usual explanatory variables included in Barro regressions, Cukierman and others consider labor market indicators like those used in the previous section.

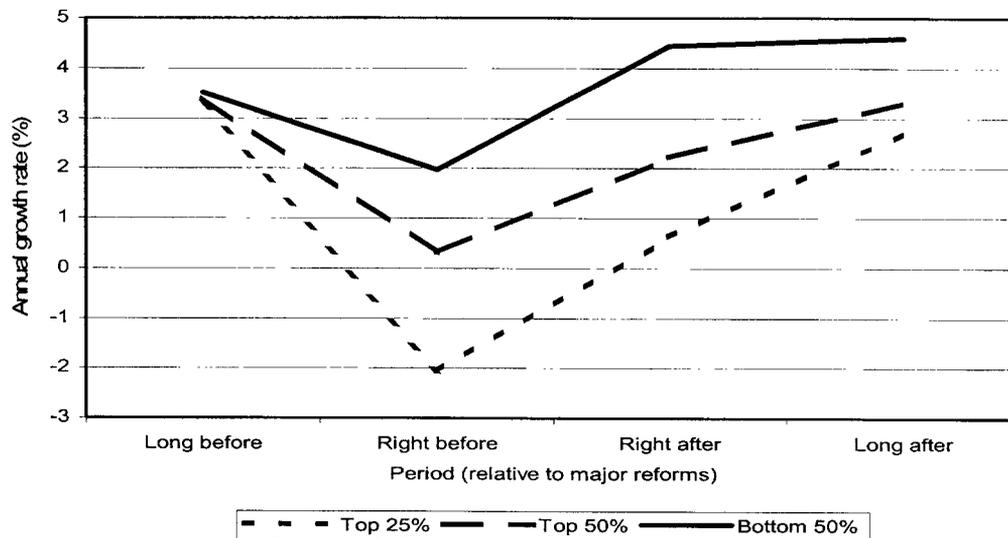
With the exception of the minimum wage, none appears to have a significant impact, positive or negative, on long-run growth. When the minimum wage is very low, an increase in this wage is associated with a higher growth rate. The effect vanishes when the minimum wage reaches roughly 20 percent of the average labor cost in manufacturing. Preliminary work on the transmission mechanism for this effect suggests two possible avenues. At a very low minimum wage, an increase is associated with higher educational attainment, a result consistent with investments in children's schooling being profitable but being constrained by lack of access to credit. An increase in the minimum wage, again from a very low level, is also associated with higher employment, a result that would make sense if employers had enough market power to set wages below their competitive equilibrium level.

The weakness of the effects of other labor market policies on long-run growth suggests that deadweight costs could be small or negligible. Such a conclusion would be consistent with the view that different labor market policies and institutions are associated with similar economic performances. In the words of Freeman (2000), the capitalist system is diversified rather than "single peaked." From this perspective, the main effect of these policies and institutions is on the distribution of labor market rents between workers and employers, not on economic efficiency. The long-run growth regressions by Cukierman and others (2001) provide some support for this view. But because of a lack of data, these growth regressions do not include separation costs among their arguments. Separation costs could in principle have greater efficiency effects. More important, an analysis of the medium-term impact of those policies and institutions on economic growth leads to a less sanguine conclusion.

A cross-country study of economic growth during periods of economic reform shows that some of the government interventions considered in the previous section can be associated with poor performance (Forteza and Rama 2001). This study compared output rates across countries with different degrees of labor market rigidity. The rigidity of labor markets was measured using an index that combines the minimum wage, the generosity of social security benefits, labor union membership, and the government share of the labor force. The comparison was run over four periods: long before (4–10 years) major economic reforms were launched, shortly before, shortly after, and long after. The timing of reforms was identified based on the accumulated volume of World Bank lending for structural and sectoral adjustment programs.

The results of this cross-country study are summarized in figure 5. The lines in this figure represent average growth rates for three groups of countries in the four phases of reform. The results are similar when controls for past performance and external shocks are taken into account. The figure confirms that labor market rigidity does not affect long-run performance, because the growth rates for all groups are very similar long before the reforms. But countries with rigid labor markets have a much worse economic performance in the years before the reforms are launched

Figure 5. Labor Market "Rigidity" and the Success of Reforms



Note: The quantiles refer to an indicator of "actual" labor market rigidity, constructed combining the level of minimum wages, the social security contribution rate, the union membership rate, and the share of the labor force employed by the central government. The higher this indicator, the more "rigid" the labor market.

Source: Constructed by the author using data from Forteza and Rama (2001).

and a much slower recovery afterward. This pattern is consistent with countries with rigid labor markets reforming less often and later, which could explain why they do worse when they reform. Interestingly, this pattern is driven not by minimum wages or mandated benefits but by unionization and, especially, by government employment.

These results are consistent with a political interpretation of the role of labor market rigidity. In developing economies, public sector employees make up a large share of union membership. Workers in protected industries also tend to be unionized. These two groups stand to lose from such reforms as trade liberalization, market deregulation, and privatization of state-owned enterprises. The more powerful these groups are, the more likely that reforms will be delayed or adopted half-heartedly. Based on this political interpretation, figure 5 suggests that the payoffs to compensating those who stand to lose from globalization can be large. It also casts doubt on the wisdom of using public sector employment as insurance against the increased economic volatility from globalization: A high level of public sector employment may hamper the adoption of economic reforms later.

Political economy considerations raise interesting issues relating to the targeting of labor market policies. The analysis in the previous section suggests that social

protection could be an effective tool of redistribution, increasing the consumption or income share of the poorest quintiles of the population. The political economy argument emerging from the analysis in this section implies that transfers should be aimed at those who stand to lose more from globalization. These could be, for example, relatively unskilled wage earners in protected sectors. Offsetting the losses of these workers may be the key to the acceptance of trade liberalization and similar reforms. But wage earners in protected sectors are not among the poorest segments of the population in developing countries.

Policies aimed at mitigating the impact of job loss can be classified in two groups: one-time interventions and permanent programs. Among one-time interventions, generous compensation and assistance to those bound to lose their jobs have often been used to reduce resistance to privatization, trade liberalization, or market deregulation in industries where insiders are vocal and powerful (Haltiwanger and Singh 1999). If perfectly designed, compensation packages for redundant workers would make them indifferent to job loss, so that economic reforms would be Pareto optimal. But experience suggests that these packages, whose design is based on rules of thumb, can lead to substantial over- or undercompensation. A more sensible basis for determining the amount of compensation needed to achieve indifference is the analysis of microeconomic data on the earnings redundant workers could hope to make after separation (Assaad 1999; Chong and Rama 2001).

Developing economies have tried a variety of more permanent programs to help workers cope with job loss. A comparative study of Latin America considered five such programs: public works, training for the unemployed, mandatory severance pay, unemployment insurance, and forced savings (World Bank 2000). Table 6 shows that these programs have vastly different costs per assisted worker. The distribution of these costs among workers, employers, and taxpayers also varies substantially. From the point of view of this article, the most interesting feature of the table is the difference among beneficiaries. Although public works and training programs for the unemployed reach the poorest population groups, forced savings reach the richest. Mandatory severance pay and unemployment insurance fall in between. These differences should be kept in mind when thinking about targeting social protection programs in the context of integration with world markets.

Conclusion

This article has addressed some questions central to the policy debate today, offering answers based on an analytical review of the literature. Its few novel results are drawn from studies under preparation, and its value added (if any) is to make sense of the microeconomic and macroeconomic evidence available. What the article does is propose a story line consistent with all the bits and pieces of knowledge that have

Table 6. Income Support Programs for the Unemployed

| Program and country | Workers legally covered by the program | Spending per beneficiary (U.S. dollars) | Cost of the program falls on | Share of beneficiaries by income or consumption quintile (%) | | | | |
|--|--|---|------------------------------|--|--------|-------|--------|---------|
| | | | | Poorest | Second | Third | Fourth | Richest |
| Public works in Argentina | All (in principle) | 3,100 | Taxpayers | 78.6 | 15.3 | 3.5 | 2.1 | 0.4 |
| Training for unemployed in Mexico | Eligible based on age and education | 393 | Taxpayers | 69.9 | 15.5 | 8.1 | 5.0 | 1.5 |
| Severance pay in Peru | Salaried, with long enough seniority | 760 | Workers and employers | 4.7 | 9.5 | 28.6 | 33.3 | 23.8 |
| Unemployment insurance in Brazil | Salaried and in social security | 664 | Workers and employers | 10.6 | 24.6 | 19.1 | 25.1 | 13.6 |
| Individual accounts in Colombia ^a | Salaried and in social security | — | Workers | 0.0 | 4.3 | n.a. | 19.1 | 76.6 |

—, Not available.

n.a., Not applicable.

^aBeneficiary data for Colombia refer to quartiles, not quintiles.

Source: Constructed by author using data from World Bank (2000).

accumulated. Reassuringly, this story line is much more nuanced than the ideological views often voiced in the policy debate. But it could nonetheless be wrong. Although all research papers conclude with the assertion that more research is needed, such a conclusion really appears to be warranted in this case. In ending, it may be useful to summarize the tentative story line emerging from this analytical survey.

Does globalization affect labor market outcomes? The answer is yes, but in uneven ways. In the short run, exposure to world markets is associated with lower wages, but foreign direct investment is associated with higher wages. Though the dispersion of wages by occupation does not seem to change much, returns to education increase with openness to trade—and they increase dramatically with foreign direct investment. In the medium run the effect of trade liberalization on wages switches from negative to positive. The effect should become even more positive in the long run, as the positive impact of openness on output per capita should lead to a sustained increase in all wages. But globalization can be a source of hardship for old, less skilled workers in protected sectors. These workers are likely to see their wages decline. The decline in their well-being could be larger than that indicated by the estimated decline in wages because trade liberalization may also be associated with higher unemployment.

Can labor market policies mitigate or offset the effects of globalization? The answer is a qualified yes. The most effective interventions are not those that increase the bargaining power of labor relative to capital, but those falling under the social protection heading. In particular, the core labor standards at the center of the policy debate today appear to be quite ineffective. Minimum wages also appear to make no significant difference. Conversely, the amount of resources channeled through the social security system (which is highly correlated with the number of ILO conventions ratified) is associated with lower overall inequality and appears to benefit even the poorest segments of the population, even if they are unlikely to be formally covered by the system.

Finally, do labor market policies reduce efficiency? The answer depends on the time horizon considered. In the long run, labor market interventions may have little impact on economic performance. If anything, minimum wages could be associated with higher growth rates, as long as they are set at moderate levels. Minimum wages and other interventions falling under the social protection heading appear to be benign in the short run because they do not interfere with the effectiveness of economic reforms. But this effectiveness is substantially reduced when special interests are powerful. High rates of union membership and especially large public sector employment are associated with poorer performance before economic reforms are adopted and a slower recovery afterward. In developing and transition economies there is a high correlation between unionization and public sector employment, and the observed regularity could simply reflect the ability of unionized workers and public sector employees to successfully oppose economic reforms.

In sum, social protection mechanisms could be an important complement for policies aimed at further integrating developing economies with the rest of the world.

These mechanisms can mitigate the increase in inequality that could result from an increased premium for skill. They can also be used to compensate those who stand to lose from globalization, defusing the opposition to reform. Social protection mechanisms can take the form of one-time interventions or permanent programs. The most appropriate mechanisms in each setting depend on the resources available and the target group. In countries where resistance to globalization is potentially strong, the compensation could be targeted to the most vocal groups, even if they are not poor. In other cases the focus could (and should) be on the poorest groups. In practice, the optimal policy would probably combine several mechanisms, such as public works and unemployment insurance, or public works and severance pay for redundant workers.

All this said, among all the government policies that could mitigate the effects of globalization on labor market outcomes, the most promising candidate may be outside the labor market. Although the impact of trade liberalization on wages can be expected to switch from negative to positive over the medium run, and presumably to highly positive in the longer run, the increase in the skill premium may not fade away so easily. The dispersion in the educational attainment of the population may thus translate into a larger dispersion in labor earnings. Reducing the dispersion in education could be the most effective way to reduce the dispersion in earnings. So although this article has dealt mainly with labor market policies, it is important to keep in mind that they may be less effective than education policies in mitigating the adverse effects of globalization.

Notes

Martín Rama is Lead Economist in the World Bank's East Asia and Pacific Region. His e-mail address is mrma@worldbank.org. An initial version of this article was written for the European online Annual World Bank Conference on Development Economics (ABCDE-Europe). Excellent research assistance was provided by Manuel Felix. Comments and suggestions by David Dollar, Alvaro Forteza, and Remco Oostendorp are gratefully acknowledged.

1. Studies include Green and others (2001) on Brazil, Beyer and others (1999) on Chile, Robbins (1997) on Colombia, Robbins and Gindling (1999) on Costa Rica, Hanson and Harrison (1999) on Mexico, and Behrman and others (2000) on Latin America as a whole.

2. This analysis is based on work in progress by Freeman and others (2001).

3. This analysis is based on work in progress by Rama and Ravallion (2001).

References

- Assaad, R. 1997. "The Effects of Public Sector Hiring and Compensation Policies on the Egyptian Labor Market." *World Bank Economic Review* 11(1):85-118.
- . 1999. "Matching Severance Payments with Worker Losses in the Egyptian Public Sector." *World Bank Economic Review* 13(1):117-54.

- Azam, J.-P., and C. Ris. 1995. "Rent-Sharing, Hold Up, and Manufacturing Sector Wages in Cote d'Ivoire." Policy Research Working Paper 2600. World Bank, Development Research Group, Washington, D.C.
- Barro, R. 1997. *Determinants of Economic Growth: A Cross-Country Empirical Study*. Cambridge, Mass.: MIT Press.
- Behrman, J. R., N. Birdsall, and M. Székely. 2000. "Economic Reform and Wage Differentials in Latin America." IADB Research Working Paper 435. Inter-American Development Bank, Washington, D.C.
- Bell, L. 1997. "The Impact of Minimum Wages in Mexico and Colombia." *Journal of Labor Economics* 15(3):S102–35.
- Beyer, H., P. Rojas, and R. Vergara. 1999. "Trade Liberalization and Wage Inequality." *Journal of Development Economics* 59(1):103–23.
- Bourguignon, F., and C. Morrisson. 1998. "The Role of Dualism." *Journal of Development Economics* 57:233–57.
- Butcher, K. F., and C. E. Rouse. 2001. "Wage Effects of Unions and Industrial Councils in South Africa." *Industrial & Labor Relations Review* 54(2):349–74.
- Chong, A., and M. Rama. 2001. "Do Compensation Packages Need to Be That Generous? Simulations for Government Employees in Guinea-Bissau." In S. Devarajan, F. H. Rogers, and L. Squire, eds., *World Bank Economists' Forum*. Washington, D.C.: World Bank.
- Cukierman, A., M. Rama, and J. van Ours. 2001. "Long-Run Growth, the Minimum Wage and Other Labor Market Institutions." Tilburg University, Center for Economic Research, Tilburg, Netherlands.
- Deininger, K., and L. Squire. 1996. "A New Data Set Measuring Income Inequality." *World Bank Economic Review* 10(3):565–91.
- Easton, S. T., and M. A. Walker. 1992. *Rating Economic Freedom*. Vancouver, B.C.: Fraser Institute.
- Fallon, P. R., and R. Lucas. 1991. "The Impact of Changes in Job Security Regulations in India and Zimbabwe." *World Bank Economic Review* 5(3):395–413.
- Feenstra, R. C., and G. H. Hanson. 1997. "Foreign Direct Investment and Relative Wages: Evidence from Mexico's Maquiladoras." *Journal of International Economics* 42(3–4):371–93.
- Forteza, A., and M. Rama. 2001. "Labor Market Rigidity and the Success of Economic Reforms across More than 100 Countries." Policy Research Working Paper 2521. World Bank, Development Research Group, Washington, D.C.
- Freeman, R. 2000. "Single-Peaked vs. Diversified Capitalism: The Relation between Economic Institutions and Outcomes." NBER Working Paper 7556. National Bureau of Economic Research, Cambridge, Mass.
- Freeman, R., and R. Oostendorp. 2000. "Wages around the World." NBER Working Paper 8058. National Bureau of Economic Research, Cambridge, Mass.
- Freeman, R., R. Oostendorp, and M. Rama. 2001. "Globalization and Wages." World Bank, Development Research Group, Washington, D.C.
- Gindling, T. H., and K. Terrell. 1995. "The Nature of Minimum Wages and Their Effectiveness as a Wage Floor in Costa Rica, 1976–91." *World Development* 23(8):1439–58.
- Green, F., A. Dickerson, and J. S. Arbache. 2001. "A Picture of Wage Inequality and the Allocation of Labour through a Period of Trade Liberalization: The Case of Brazil." *World Development* 29(11):1923–39.
- Gruber, J. 1997. "The Incidence of Payroll Taxation: Evidence from Chile." *Journal of Labor Economics* 15(3):S72–101.
- Haltiwanger, J., and M. Singh. 1999. "Cross-Country Evidence on Public Sector Retrenchment." *World Bank Economic Review* 13(1):23–66.

- Hanson, G. H., and A. Harrison. 1999. "Trade Liberalization and Wage Inequality in Mexico." *Industrial & Labor Relations Review* 52(2):271–88.
- Heckman, J., and C. Pagés. 2000. "The Cost of Job Security Regulation: Evidence from Latin American Labor Markets." NBER Working Paper 7773. National Bureau of Economic Research, Cambridge, Mass.
- Heston, A., R. Summers, and B. Aden. 2002. *Penn World Table*, version 6.1. Available online at <http://pwt.econ.upenn.edu/aboutpwt.html>.
- Li, H., L. Squire, and H. Zou. 1998. "Explaining International and Intertemporal Variations in Income Inequality." *Economic Journal* 108:26–43.
- MacIsaac, D., and M. Rama. 1997. "Determinants of Hourly Earnings in Ecuador: The Role of Labor Market Regulations." *Journal of Labor Economics* 15(3):S136–65.
- . 2001. "Mandatory Severance Pay: An Assessment of Its Coverage and Effects in Peru." Policy Research Working Paper 2626. World Bank, Development Research Group, Washington, D.C.
- Maloney, W. F., and J. Nuñez. 2003. "Measuring the Impact of Minimum Wages: Evidence from Latin America." NBER Working Paper 9800. National Bureau of Economic Research, Cambridge, Mass.
- Morrison, A. R. 1994. "Are Institutions or Economic Rents Responsible for Inter-industry Wage Differentials?" *World Development* 22:355–68.
- Mwabu, G., and T. P. Schultz. 2000. "Wage Premiums for Education and Location of South African Workers, by Gender and Race." *Economic Development and Cultural Change* 48(2):307–34.
- Nelson, J. 1991. "Organized Labor, Politics, and Labor Market Flexibility in Developing Countries." *World Bank Research Observer* 6(1):37–56.
- Panagides, A., and H. A. Patrinos. 1994. "Wage Differentials in the Developing World: A Case Study of Mexico." Policy Research Working Paper 1269. World Bank, Latin American and Caribbean Country Department, Washington, D.C.
- Psacharopoulos, G. 1994. "Returns to Investment in Education: A Global Update." *World Development* 22:1325–43.
- Rama, M. 2000. "Wage Misalignment in CFA Countries: Are Labor Market Policies to Blame?" *Journal of African Economies* 9(4):475–511.
- . 2001. "The Consequences of Doubling the Minimum Wage: The Case of Indonesia." *Industrial & Labor Relations Review* 54(4):864–81.
- . 2003. "The Sri Lankan Unemployment Problem Revisited." *Review of Development Economics* 7(3):510–25.
- Rama, M., and R. Artecona. 2002. "A Database of Labor Market Indicators across Countries." World Bank, Development Research Group, Washington, D.C.
- Rama, M., and M. Ravallion. 2001. "Labor Market Regulation and Inequality: A Cross-Country Analysis." World Bank, Development Research Group, Washington, D.C.
- Robbins, D. 1997. "Trade and Wages in Colombia." *Estudios de Economía* 24(1):47–83.
- Robbins, D., and T. H. Gindling. 1999. "Trade Liberalization and the Relative Wages for More-Skilled Workers in Costa Rica." *Review of Development Economics* 3(2):140–54.
- Rodrik, D. 2000. "What Drives Public Employment in Developing Countries?" *Review of Development Economics* 4(4):229–43.
- . 1999. "Democracies Pay Higher Wages." *Quarterly Journal of Economics* 114(3):707–38.
- Sachs, J. D., and A. M. Warner. 1995. "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity* 1. Washington, D.C.: Brookings Institution.
- Sala-i-Martin, X. 1997. "I Just Ran Two Million Regressions." *American Economic Review, Papers and Proceedings* 87(2):178–83.

- Terrell, K., and J. Svejnar. 1989. *The Industrial Labor Market and Economic Performance in Senegal*. Boulder, Colo.: Westview Press.
- Wood, A. 1994. *North-South Trade, Employment and Inequality*. Oxford: Clarendon Press.
- World Bank. 1990. *World Development Report 1990: Poverty*. New York: Oxford University Press.
- . 1993. *The East Asian Miracle: Economic Growth and Public Policy*. World Bank Policy Research Report. New York: Oxford University Press.
- . 2000. *Securing Our Future in a Global Economy*. Latin American and Caribbean Studies Series. Washington, D.C.
- . 2001 *Globalization, Growth and Poverty*. World Bank Policy Research Report. New York: Oxford University Press.

Developing Countries' Changing Participation in World Trade

Will Martin

Recent years have seen substantial reductions in trade policy and other barriers inhibiting developing country participation in world trade. Lower barriers have contributed to a dramatic shift in the pattern of developing country trade—away from dependence on commodity exports to much greater reliance on manufactures and services. In addition, exports to other developing countries have become much more important. These changes have profound implications for the role played by developing countries in the world economy and trade system.

Over the past 50 years, trade's role in the world economy has increased dramatically. Trade has grown much faster than output, and most of the countries that have achieved the fastest economic growth have done so by rapidly increasing their participation in world trade. Import substitution policies—widely used in the 1950s, 1960s, and 1970s—were much less successful than the export-oriented policies used in East Asia's high-growth economies. As a result, policymakers in many developing economies began adopting policies for more open trade during the 1980s. By the late 1980s, nearly all of the centrally planned regimes that had eschewed market-based trade had either collapsed or made dramatic reforms that gave foreign trade and investment a prominent place in their development programs.

These reforms led to dramatic changes in developing economies' involvement in international trade. Until the mid-1980s, most developing areas relied primarily on commodity exports, which exposed them to the high volatility of and continuous decline in commodity prices and created concerns about dependence on imported manufactures. But in the early 1980s many developing economies began to dramatically increase manufactured exports. By the late 1990s about 80 percent of exports from developing economies were manufactured goods, greatly easing earlier concerns about the role of trade. Thus poor countries, home to some 3 billion

people, are now active in global markets for manufactures and services (World Bank 2002a).

But continued increases in integration are far from inevitable—as is clear from the concerns about globalization manifested in huge protests in virtually every city that has hosted a major international policy meeting since 1999. These concerns, and the enormous costs and risks associated with international terrorism, pose serious challenges to closer economic relations between countries. Integrating with the world is a choice that must be made by policymakers, though the costs of withdrawing are considerable (Wolf 2001). Moreover, there is strong interdependence between the decisions of policymakers. If some major countries turn away from world markets, as in the 1930s, the result could be a downward spiral in international trade that hurts even those countries that want to remain integrated with the world economy.

Choosing policies to manage interactions with the world economy is not simple. Trade policy reform increasingly requires more than just reducing and streamlining border barriers, it also requires developing institutions. Many of the “behind-the-border” reforms needed to take advantage of the opportunities created by open trade regimes require institutional capacity that is in short supply in developing economies. In addition, supporting policies are likely to be needed to alleviate adverse impacts on particular groups. Globalization will likely be sustainable only if it is accompanied by policies that equip people to take advantage of the benefits offered by increased integration with the world economy (Rodrik 1997).

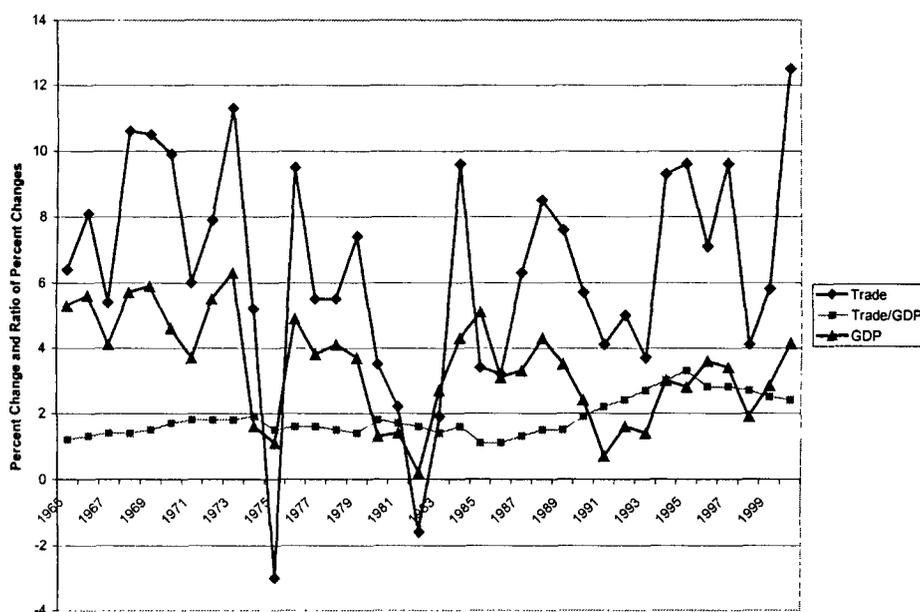
To adequately answer the question of what should be done about trade policies, it is first necessary to examine the current situation and the scope for improvement. Thus the next section examines recent changes in trade policies and other potential trade barriers. After that the article examines some major changes in the trade patterns of developing countries.

I. Recent Changes in Trade Policies and Other Barriers to Integration

Since the mid-1950s there has been extraordinary growth in world trade and economic openness. Since 1965 world trade has grown faster than world income in all but a few years of cyclical downturns (figure 1). In the 1990s trade grew more than twice as fast as income.

This rapid growth in the openness of world economies reflects a number of factors, including reductions in trade barriers, transport costs, and communication costs. Related influences include the increasing importance of trade in manufactures, for which two-way trade is far more prevalent than for commodities, and the fragmentation

Figure 1. World Trade Growth, GDP Growth, and Smoothed Trade-GDP Growth, 1965–2000 (%)



Source: World Bank (2001).

of production processes, which leads to much more international trade in components and services (Ng and Yeats 1999; Deardorff 2001).

In industrial countries, reductions in protection from the high levels reached in the 1930s were already under way when the General Agreement on Tariffs and Trade (GATT) was established in 1947, and they continued under successive rounds of GATT negotiations. Reforms involved reducing tariffs and abolishing quantitative restrictions such as those introduced for balance of payments reasons.

But the liberalization process was neither smooth nor continuous. In the 1950s, agriculture escaped most multilateral disciplines. In the 1960s and 1970s, exports of textiles and clothing from developing economies were subjected to a system of quotas that discriminated by country and violated all the fundamental principles of the GATT. As tariffs fell, nontariff barriers—such as voluntary export restraints and antidumping rules—became more important.

Even so, for the manufactures trade at the heart of the GATT liberalization process, reductions in protection were dramatic in industrial countries. From rates of 50 percent or higher in the late 1940s, average tariffs on imports of manufactures had fallen to 4.1 percent in 1988, early in the Uruguay Round (Abreu 1996). Nontariff barriers also fell for most of these goods. Trade in manufactures between industrial countries grew rapidly, much of it involving two-way trade in products within the same type of good.

In developing countries trade liberalization occurred much more slowly. The belief that infant industries could be nurtured behind protective barriers was appealing, and taxes on commodity exports were a convenient way to raise government revenues. The dependency concerns enunciated by Raoul Prebisch provided intellectual support for these approaches by associating open markets with dependence on commodities and identifying declines in commodity prices with declines in the terms of trade for such countries. In addition, import substitution policies were strongly supported by the vested interests that gained from protection. Tariff barriers became ubiquitous, as did nontariff barriers such as quotas and licenses. Foreign exchange restrictions often imposed large additional taxes on trade.

Under import substitution regimes, import quotas frequently generated quota rents equal to large percentages of gross domestic product (GDP) (Krueger 1980). Quotas and licenses were often allocated in a way that provided strong incentives to use real resources in pursuit of the rents associated with them, resulting in wasteful dissipation of scarce resources (Bhagwati 1983). Such incentives for rent seeking can have extremely negative consequences for development. In many developing economies, these incentives appear to have driven the social returns on massive investments in human capital to zero or negative levels (Pritchett 2001). They also encouraged firms to locate near capital cities to provide better access to political decision-makers, rather than in areas where environmental costs would be lower or production more efficient (Livas Elizondo and Krugman 1992). Finally, such policies encourage corruption, with its sharply adverse consequences for institutional quality, development, and equity (World Bank 2000).

The protective effects of tariffs, quotas, and licenses in developing countries were often reinforced by distortions in foreign exchange markets. These distortions involved an overvalued official exchange rate coupled with some mechanism to allocate scarce foreign exchange among uses. Where such allocations are made through legal secondary markets, as in China between the early 1980s and 1994 (Martin 1993b) or Nigeria in recent years, the resulting trade distortions are reduced because scarce foreign exchange is allocated to those who place the highest value on it but can be meaningfully measured and are frequently very substantial.

With the profoundly distorted trade and foreign exchange regimes in most developing economies prior to the 1980s, it was difficult to know which policy reforms would improve welfare. Some limited guidelines were available. For instance, policy changes that increased transparency and unified distortions across firms generally improved welfare—as when foreign exchange allocation was replaced by a secondary market approach (see Dervis, de Melo, and Robinson 1981 for an example). But simply opening up to foreign investment, without reducing barriers, could easily reduce welfare (Brecher and Diaz-Alejandro 1977). Policies that imposed export performance requirements on foreign investors could be second-best welfare improving (Rodrik 1987).

The political attractions of import substitution policies for developing-country policymakers were not effectively countered by the multilateral trade system prior to the Uruguay Round of GATT negotiations (1986–94). Developing-country policymakers generally focused their efforts in the GATT on obtaining unreciprocated improvements in access to industrial country markets under the rubric of “special and differential treatment.” This approach may have helped some beneficiary countries by improving their terms of trade.¹ But it also had adverse economic consequences.

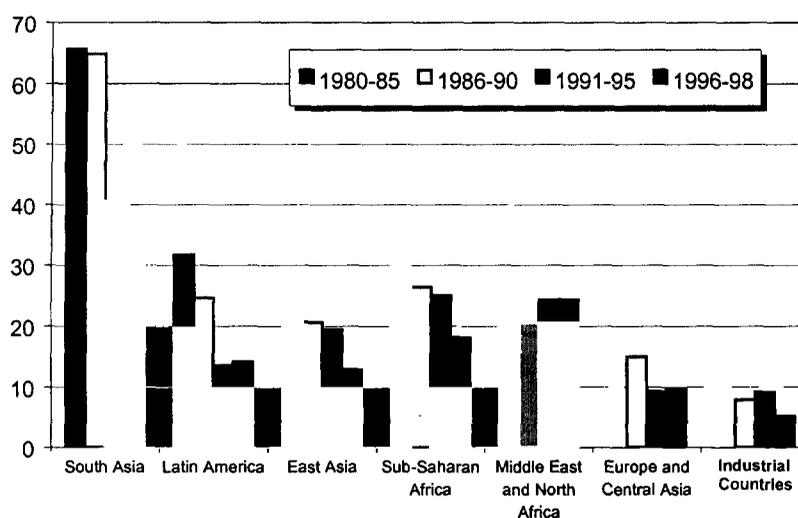
First, the approach made it difficult for developing economies to bargain for improvements in market access in the products of greatest interest to them. Second, such access was often constrained by quantitative restrictions and the risk of preference erosion or removal. Third, this approach meant that domestic exporters had no incentive to lobby for reductions in domestic protection as part of a package to improve their access to partner markets. It was no coincidence that industrial countries introduced new barriers in areas of particular interest to developing countries—such as agriculture and textiles and clothing—during this period.

A small group of developing economies, primarily in East Asia, either did not pursue import-substitution policies or used complementary export promotion policies as part of an export-oriented development path. By the late 1960s, the outstanding performance of economies such as Hong Kong, the Republic of Korea, and Taiwan (China) had begun to attract attention. This performance, combined with critical assessments of import substitution regimes (see Krueger 1980), contributed to a gradual evolution in thinking toward more outward-oriented policies.

In light of this experience, developing countries began to adopt more open trade regimes. This change in attitude first manifested itself at the multilateral level in the Uruguay Round, when developing countries began to play the main game of exchanging market access concessions. Developing countries agreed to bind their tariffs—that is, commit to not increase them above specified levels recorded in a schedule of concessions at the World Trade Organization—on 100 percent of their agricultural imports and on more than 60 percent of their imports of industrial products (Abreu 1996).

But the most profound and far-reaching manifestation of developing countries' interest in greater participation in trade comes from the wave of unilateral trade reforms that has swept developing countries. These reforms have affected all regions and all the main types of policy distortions. Between the early 1980s and late 1990s average tariffs in developing countries were cut in half (figure 2). By 1999, average tariffs in developing countries were just 11 percent (see www.worldbank.org/trade). Absolute reductions in tariffs in developing countries have been much higher than in industrial countries, and decreases from higher levels are likely to have a much greater welfare benefit than corresponding decreases from lower levels (see Martin 1997). In addition, the dispersion of tariff rates, which typically increases the welfare cost of any average rate, was substantially reduced.

Figure 2. Average Tariffs by Region, 1980–98 (%)



Source: World Bank (2001).

One must be careful when examining changes in tariffs, because a decline may reflect substitution of nontariff barriers for tariffs. But during this period nontariff barriers in developing countries also fell considerably (table 1). Another important dimension of reform has been the sharp reduction in the number of developing countries using foreign exchange restrictions and in average foreign exchange premiums. In the 1980s foreign exchange market distortions were enormous, but these premiums have since fallen to very low levels in most developing countries and regions (table 2). At 46.5 percent, the Middle East and North Africa has the highest average foreign exchange rate premium (calculated as a simple average), but this is

Table 1. Core Nontariff Barriers in Developing Regions, 1989–98 (%)

| Region | 1989–94 | 1995–98 |
|--------------------------------------|---------|---------|
| East Asia and the Pacific (7) | 30.1 | 16.3 |
| Latin America and the Caribbean (13) | 18.3 | 8.0 |
| Middle East and North Africa (4) | 43.8 | 16.6 |
| South Asia (4) | 57.0 | 58.3 |
| Sub-Saharan Africa (12) | 26.0 | 10.4 |

Note: Numbers in parentheses are the number of countries in each region for which data are available.

Source: World Bank (2001) based on Michalopoulos (1999).

Table 2. Average Black Market Premiums in Developing Regions, 1980–97 (%)

| <i>Region</i> | <i>1980–89</i> | <i>1990–93</i> | <i>1994–97</i> |
|----------------------------------|----------------|----------------|----------------|
| Total ^a | 82.0 | 78.2 | 20.3 |
| East Asia | 3.6 | 3.6 | 3.2 |
| Latin American and the Caribbean | 48.7 | 13.1 | 4.4 |
| Middle East and North Africa | 165.6 | 351.6 | 46.5 |
| Excluding Algeria and Iran | 7.1 | 8.8 | 1.4 |
| South Asia | 40.8 | 45.1 | 10.1 |
| Sub-Saharan Africa | 116.5 | 28.6 | 32.2 |
| Excluding Nigeria | 112.1 | 25.8 | 9.6 |

^aSample of 41 developing countries.

Source: World Bank (2001).

almost entirely due to Algeria and Iran. When these two outliers are excluded, the average rate falls to just 1.4 percent. Similarly, when Nigeria is excluded the average premium in Sub-Saharan Africa is less than 10 percent, down from 112 percent in the mid-1980s.

A detailed study on trade regimes in Africa confirms that the region's trade reforms in the 1990s were substantial and wide ranging, including reductions in tariffs and elimination of foreign exchange distortions and export taxes and monopolies (Hinkle and Herrou-Aragon 2002). Key remaining areas of concern include corruption and inefficiency in customs administration, as well as problems in ensuring that exporters have access to intermediate inputs at close to world prices.

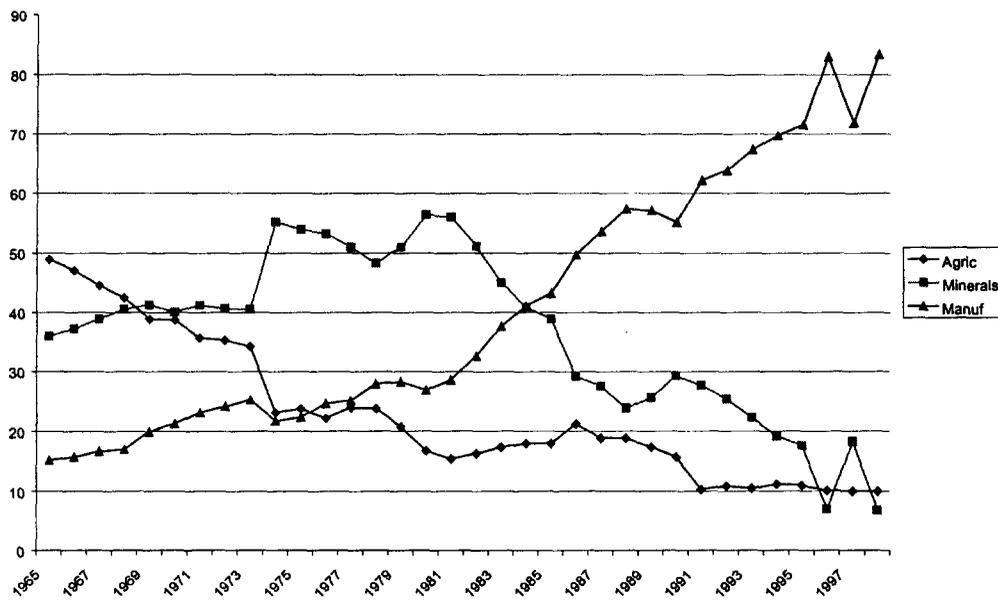
II. Changes in Trade Patterns

Changes in trade policies and reductions in trade barriers have been associated with major changes in developing countries' role in the world economy. In particular, as developing economies have lowered their trade barriers, the composition of their exports has changed enormously. Since the 1980s, developing countries have drastically increased exports of manufactures and exports to other developing areas. In addition, exports of services have become much more important. This section first considers developments in the exports of developing countries as a whole, then changes in the exports of selected groups of relatively commodity-dependent economies.

Exports of All Developing Countries

In 1965 agricultural commodities accounted for about 50 percent of developing country exports, and manufactures for only about 15 percent (figure 3). Since then

Figure 3. Composition of Merchandise Exports from Developing Countries, 1965–98 (% total)



Source: GTAP 5 database.

the share of manufactures has risen continuously, except for a transient decline in 1997 associated with the East Asian crisis. The share of agricultural products has shown a similarly consistent decline, to about 10 percent by 1998.

The share of metals and minerals has fluctuated, with a rise in 1973 associated with higher oil prices among members of the Organization of Petroleum Exporting Countries, lifting the share of this category from 1973 through the early 1980s. But since the early 1980s, the share of mineral exports has declined, falling below 10 percent by 1998. Overall, developing countries have been leading the change in the pattern of world trade, transforming their export pattern from commodities to manufactures much more rapidly than the world as a whole.

These dramatic changes in developing country exports have many underlying causes. One is the relatively high rates of accumulation of human and physical capital in these countries. Estimates by Nehru and Dhareshwar (1993) suggest that between 1960 and 1990 the ratio of capital to labor more than doubled in most developing countries. Estimates by Nehru, Swanson, and Dubey (1995) suggest that education per worker, particularly secondary education per worker, rose even faster. Although such estimates need to be treated with caution (Pritchett 2000, 2001), trade theory suggests that rapid growth in these factors should increase both production and trade in sectors that use them intensively. Gehlhar, Hertel, and

Martin (1994) find that rapid accumulation of these factors contributed to a strong shift out of agricultural activities and into export-oriented manufacturing activities in East Asian economies.

The highly protectionist policies followed by most developing countries prior to the 1980s were often designed at least partly to stimulate industrialization. But one of their effects was to greatly constrain countries' ability to participate in the more dynamic parts of international trade—trade in manufactures and services. Both of these typically require access to capital, technology, and intermediate inputs that are often best obtained from abroad. Martin (2003) shows that the cost structure and pattern of protection in developing countries discriminated heavily against manufactures and agricultural processing, contributing to dependence on agricultural and resource exports. Even if high protection or high transport costs did stimulate industrialization in developing countries at an earlier stage, it appears that lowering barriers now tends to do so (Krugman and Venables 1995). Amiti (2003) points out that liberalizing more capital-intensive activities producing intermediate goods may help labor-abundant countries develop the agglomerations of industry that are a key feature of industrialization.

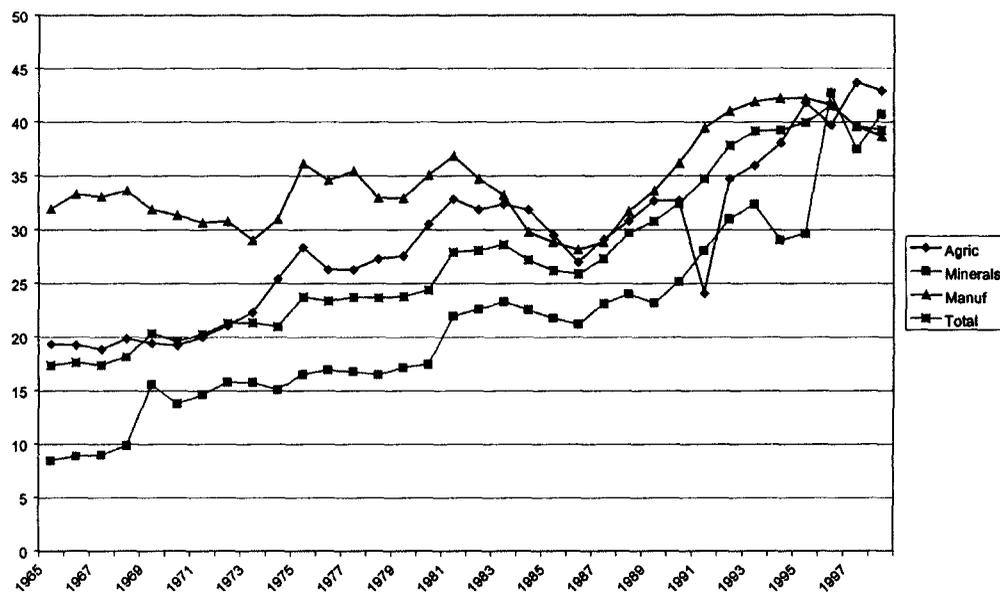
The shift away from agriculture in developing areas does not appear to have been driven by higher productivity growth in manufacturing. In fact, in recent decades productivity growth has been higher for agriculture in both industrial and developing countries (Martin and Mitra 2001; Bernard and Jones 1996). Nor have transport costs fallen as rapidly and consistently as many suggest. Hummels (1999) finds that liner shipping charges actually rose by more than 50 percent in real terms between 1954 and 1983 before beginning to decline. But transport became much faster and more reliable, and the costs of air transport declined very quickly—which Hummels (2001) estimates was equivalent to more than a 20 percent point reduction in tariffs.

Along with changes in the commodity composition of exports have come substantial changes in their direction. In the mid-1960s only 17 percent of developing country exports were destined for other developing countries (figure 4). By 1995 that share had increased to more than 40 percent. Although this share fell in 1996–98, even in 1998 it was twice as high as in the mid-1960s.

Interestingly, the increase in the share of exports going to other developing countries is not just due to the increase in manufactured exports. In the 1960s the share of manufactured exports going to other developing countries was much higher than the share of agricultural and mineral exports, which went primarily to industrial countries. But shares of agricultural and mineral products exported to other developing countries have risen substantially over time.

This increase in the importance of developing countries as markets for each other's goods reflects several factors, including growth in the share of developing countries in world trade and liberalization of developing country trade. Whalley

Figure 4. Developing Country Exports to Other Developing Countries, 1965–98 (% total exports)



Source: GTAP 5 database.

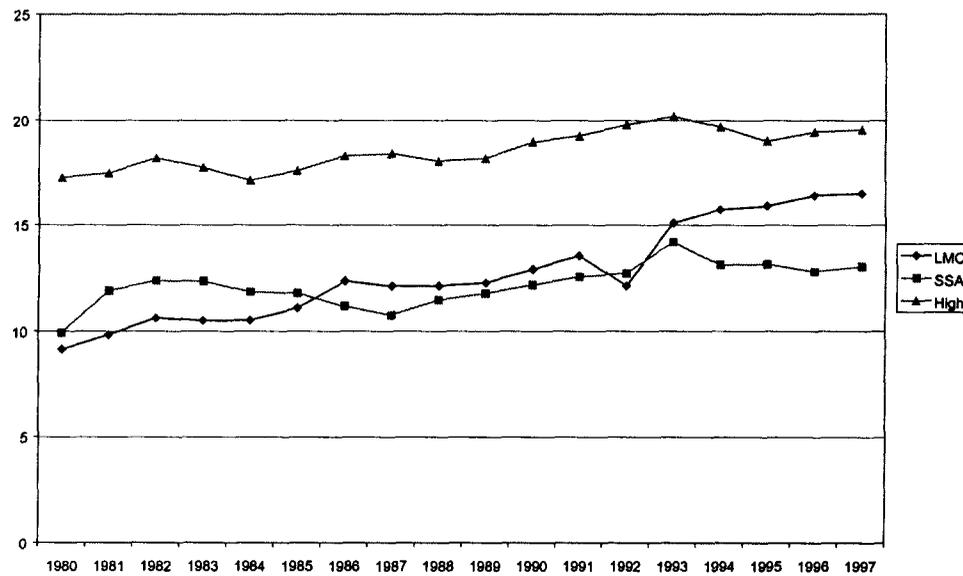
(1985) predicted that unilateral, nondiscriminatory liberalization of the type observed in developing countries in recent decades (World Bank 2001) would substantially increase trade between these countries. This increase in the importance of developing country markets has important implications for the market access barriers facing developing countries. Today most of the tariffs they face on exports are imposed by other developing countries (Hertel and Martin 2001).

Another important change in world trade has been a substantial increase in the importance of trade in services. In the early 1980s commercial services made up 17 percent of the exports of high-income countries—a share that has since risen to 20 percent (figure 5). In developing economies, trade in these services started out much less important, at 9 percent, but rose much more rapidly, to 17 percent. Even in Sub-Saharan Africa this share grew rapidly, from 10 percent to 15 percent.

Exports of Certain Groups of Developing Countries

The considerable success of developing countries in transforming their export patterns may be consistent with a group of these countries being left behind, remaining dependent on traditional commodity exports. To see whether major groups of developing countries have been left out, it is useful to examine developments in

Figure 5. Exports of Commercial Services from Developing Countries, 1980–97 (% total exports)



Note: The services exports represented here are somewhat narrower than the concept of trade in services used in the General Agreement on Trade in Services (GATS). Commercial services is a balance of payments concept covering services traded across borders (GATS mode 1) or through movement of the consumer (GATS mode 2). It excludes services traded by establishing a service-providing firm in the consuming country (GATS mode 3) or by temporary movement of a service provider (GATS mode 4).

Source: World Bank (2001, 2002b.)

individual countries and groups. The Appendix presents information on the shares of manufactures in developing country exports for the 65 countries covered by the Global Trade Analysis Project (GTAP) 5 database for 1997.

The Appendix shows the very high share of manufactures in the exports of a wide range of developing economies. Many developing countries—including low-income countries, such as Bangladesh, China, and Sri Lanka—have shares above the world average of 81 percent. Others, such as India, Indonesia, Morocco, and Turkey, have shares nearly as high as the world average.

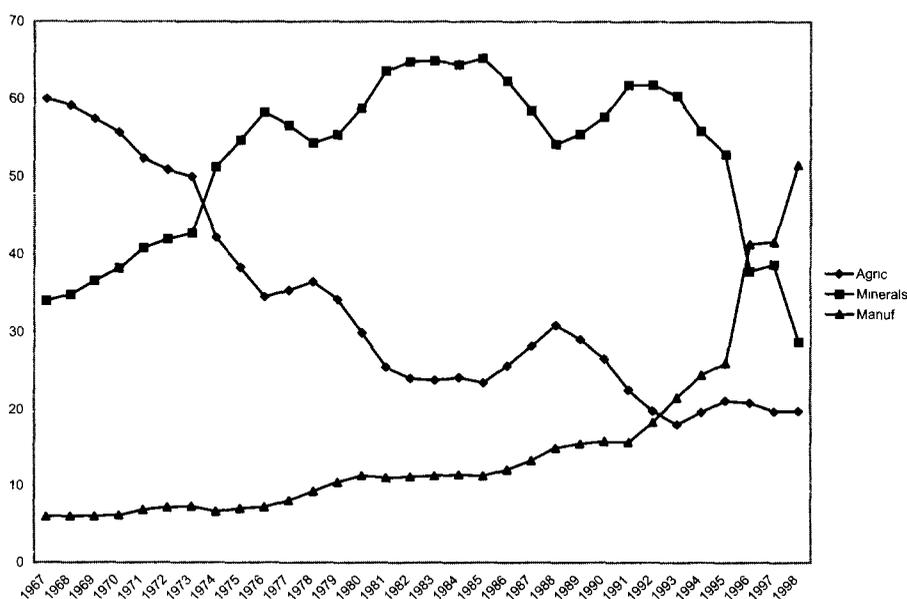
Countries where the share of manufactures remains low include reform-oriented countries, such as Chile and Uganda, and a large number of countries in Sub-Saharan Africa, the Andean region, and the Middle East and North Africa that have been less successful in integrating with the world economy. Clearly a lot depends on the composition of exports in these commodity groups. If countries have been developing new commodity exports and markets, as in Chile, then heavy dependence on commodity exports may be consistent with rapid growth in exports.

But for regions such as Sub-Saharan Africa, there is more of a concern that continuing dependence on commodity exports reflects a lack of investment in the infrastructure, equipment, and skills needed for successful participation in global trade. Still, although shares of agricultural and mineral exports from Sub-Saharan Africa have varied because of problems associated with missing data for some important countries and years, there has been a consistent increase in the share of manufactured exports from the region—from 10 percent in the mid-1960s to about 25 percent in the late 1990s (figure 6). Although low relative to others, change has been substantial even in this region.

Obviously the results for a large region such as Sub-Saharan Africa may be influenced by changes in a few large economies. To guard against this problem, data on trade patterns were analyzed for five of the region's poorer countries: Malawi, Mozambique, Tanzania, Zambia, and Zimbabwe. Even in this group the share of manufactures in total exports has increased substantially, from 2 percent in the mid-1960s to roughly 25 percent in recent years.

Together, the striking changes in developing country export patterns indicate a fundamental shift away from the traditional North-South model of the world

Figure 6. Composition of Exports from Sub-Saharan Africa, 1965–88 (% total)



Source: GTAP 5 database.

economy, in which developing economies export commodities in return for imports of manufactures. This diversification of exports and shift away from commodities offers many advantages. In particular, it reduces the volatility of export returns and diminishes concerns about potential price declines as exports expand (Martin 1993a). This is especially the case for developing countries that can promote shifts to more capital- and technology-intensive exports by promoting capital accumulation and increasing the skills of their workers (Mayer 2001).

But the dramatic increases in manufactured exports from developing countries have also contributed to protectionist concerns in both industrial and developing countries and to the emergence of new concerns about issues such as labor standards. Thus these changes have important implications for trade policy and for broader thinking about development. In any event, the frequently cited stylized fact (see, for example, Todaro 1994:52) that developing countries depend on primary product exports surely requires reexamination.

III. Conclusion

The world has been rapidly integrating in recent decades. Although this process appeared to have broad support in the early to mid-1990s, it has recently come under sustained attack. Whether to continue the process has become a pressing policy question—one with enormous implications.

An important feature of the integration process has been the major shift in trade relations between industrial and developing countries. This shift, from dependence on exports of commodities to much greater reliance on exports of manufactures and services, is sufficiently pronounced as to require rethinking of old views of trade and development. It has also placed major pressures for change on the multilateral trade system because developing countries have become much more active participants. Of course, this shift has not affected all countries equally, and many countries—particularly those in conflict—have failed to participate. But for countries that have participated, this change has profound implications. Not only does it greatly diminish concerns about potential declines in the terms of trade, it also puts much greater pressure on policymakers to maintain a relatively open regime—one that allows imports of intermediate and capital goods and supports production of manufactured goods for exports.

**Appendix. Shares Of Manufactures In Total
Merchandise Exports, Various Economies (%)**

| <i>Economy</i> | <i>Share</i> |
|--------------------|--------------|
| Japan | 98.1 |
| Taiwan, China | 96.3 |
| Singapore | 96.0 |
| Hong Kong, China | 95.9 |
| Korea, Rep. of | 94.4 |
| Sweden | 94.3 |
| Finland | 92.7 |
| Austria | 92.2 |
| Italy | 92.2 |
| Germany | 92.1 |
| Portugal | 91.4 |
| China | 90.7 |
| Philippines | 89.3 |
| Bangladesh | 89.3 |
| Switzerland | 87.5 |
| United States | 86.8 |
| Belgium/Luxembourg | 86.3 |
| Sri Lanka | 85.9 |
| United Kingdom | 85.3 |
| France | 84.2 |
| Malaysia | 84.1 |
| Thailand | 83.2 |
| Hungary | 83.2 |
| Mexico | 81.7 |
| Ireland | 81.6 |
| Spain | 81.4 |
| World average | 81.2 |
| Turkey | 79.8 |
| Poland | 77.4 |
| India | 76.8 |
| Canada | 76.5 |
| Netherlands | 73.6 |
| Morocco | 69.3 |
| Denmark | 68.4 |
| Indonesia | 62.1 |
| Greece | 61.8 |
| Brazil | 59.1 |
| Vietnam | 56.8 |
| Uruguay | 47.4 |
| Argentina | 39.7 |
| Venezuela | 37.7 |
| New Zealand | 36.3 |
| Colombia | 33.6 |

Appendix. (continued)

| | |
|------------|------|
| Australia | 32.1 |
| Zimbabwe | 31.2 |
| Chile | 24.3 |
| Mozambique | 19.6 |
| Peru | 18.3 |
| Tanzania | 11.8 |
| Malawi | 9.8 |
| Zambia | 9.7 |
| Uganda | 1.2 |

Source: GTAP 5 database.

Notes

Will Martin is a lead economist in the Development Research Group at the World Bank. His email address is wmartin1@worldbank.org.

1. In an article that sharply challenges the general view, Rose (2002) concludes that membership in the World Trade Organization (wto) has not been associated with greater liberalization or trade, whereas access to trade preferences has substantially increased trade. This result may well be driven by the large number of developing countries that used special and differential provisions to avoid making substantial reductions in protection. If found to apply more generally, Rose's conclusions would require considerable rethinking of the determinants of trade liberalization and trade growth.

References

- Abreu, Marcelo. 1996. "Trade in Manufactures: The Outcome of the Uruguay Round and Developing Country Interests." In Will Martin and L. Alan Winters, eds., *The Uruguay Round and the Developing Countries*. Cambridge: Cambridge University Press.
- Amiti, Mary. 2003. "How the Sequence of Trade Liberalization Affects Industrial Location." Paper presented to the Festschrift conference for Professor Peter Lloyd, University of Melbourne, 23–24 January. Available online at <http://pluto.ecom.unimelb.edu.au/conference2002>.
- Bernard, Andrew and Charles Jones. 1996. "Productivity across Industries and Countries: Time Series Theory and Evidence." *Review of Economics and Statistics* 78(1):135–46.
- Bhagwati, Jagdish. 1983. "DUP Activities and Rent Seeking." *Kyklos* 36(4):634–37.
- Brecher, R. and Carlos Diaz-Alejandro. 1977. "Tariffs, Foreign Capital, and Immiserizing Growth." *Journal of International Economics* 7:317–22.
- Deardorff, Alan. 2001. "International Provision of Trade Services, Trade, and Fragmentation." *Review of International Economics* 9(2):233–48.
- Dervis, Kemal, Jaime de Melo, and Sherman Robinson. 1981. "General Equilibrium Analysis of Foreign Exchange Shortages in a Developing Country." *Economic Journal* 91(December):891–906.
- Gehlhar, Mark, Thomas Hertel, and Will Martin. 1994. "Economic Growth and the Changing Structure of Trade and Production in the Pacific Rim." *American Journal of Agricultural Economics* 76(December):1101–10.
- Hertel, Thomas and Will Martin. 2001. "Liberalizing Agriculture and Manufactures in a Millennium Round: Implications for Developing Countries." In B. Hoekman and W. Martin, eds., *Developing Countries and the WTO: A Pro-Active Agenda*. Oxford: Blackwell.
- Hinkle, Lawrence and A. Herrou-Aragon. 2002. *How Far Did Africa's First Generation Trade Reforms Go?* Vol. 1. Washington, D.C.: World Bank.
- Hummels, David. 1999. "Have International Transport Costs Declined?" Purdue University, West Lafayette, Ind.
- . 2001. "Time as a Trade Barrier." Purdue University, West Lafayette, Ind.
- Krueger, A. O. 1980. "Trade Policy as an Input to Development." *American Economic Review* 70(2):288–92.
- Krugman, Paul and Anthony Venables. 1995. "Globalization and the Inequality of Nations." *Quarterly Journal of Economics* 110(4):857–80.
- Livas Elizondo, Raul and Paul Krugman. "Trade Policy and the Third World Metropolis." NBER Working Paper 4283. National Bureau of Economic Research. Cambridge, Mass.

- Martin, Will. 1993a. "The Fallacy of Composition and Developing Country Exports of Manufactures." *World Economy* 16(2):159-72.
- . 1993b. "Modeling the Post Reform Chinese Economy." *Journal of Policy Modeling* 15(5&6):545-79.
- . 1997. "Measuring Welfare Changes with Distortions." In J. Francois and K. Reinert, eds., *Applied Methods for Trade Policy Analysis*. Cambridge: Cambridge University Press.
- . 2003. "Outgrowing Resource Dependence." World Bank, Washington, D.C.
- Martin, Will and Devashish Mitra. 2001. "Productivity Growth and Convergence in Agriculture versus Manufacturing." *Economic Development and Cultural Change* 49(2):403-22.
- Mayer, Jorge. 2001. "The Fallacy of Composition: A Review of the Literature." Paper presented at the United Nations Conference on Trade and Development, Geneva.
- Michalopoulos, Costas. 1999. "Trade Policy and Market Access Issues for Developing Countries." Policy Research Working Paper 2214. World Bank, Washington, D.C.
- Nehru, Vikram., and Ashok Dhareshwar. 1993. "A New Database on Physical Capital Stock: Sources, Methodology and Results." *Revista de Analisis Economico* 8(1):37-59.
- Nehru, Vikram, Eric Swanson, and Ashutosh Dubey. 1995. "A New Database on Human Capital Stock: Sources, Methodology and Results." *Journal of Development Economics* 46(April):379-401.
- Ng, Francis and Alexander Yeats. 1999. "Production Sharing in East Asia: Who Does What for Whom, and Why?" Policy Research Working Paper 2197. World Bank, Washington, D.C.
- Pritchett, Lant. 2000. "The Tyranny of Concepts: CUDIE (Cumulated, Depreciated Investment Effort) Is Not Capital." *Journal of Economic Growth* 5(4):361-84.
- . 2001. "Where Has All the Education Gone?" *World Bank Economic Review* 15(3):367-91.
- Rodrik, Dani. 1987. "The Economics of Export Performance Requirements." *Quarterly Journal of Economics* 102(August):633-50.
- . 1997. *Has Globalization Gone Too Far?* Washington, D.C.: Institute for International Economics.
- Rose, Andrew. 2002. "Do We Really Know That the WTO Increases Trade?" NBER Working Paper 9273. National Bureau of Economic Research, Cambridge, Mass.
- Todaro, M. 1994. *Economic Development*. 5th ed. New York: Longman.
- Whalley, John 1985. *Trade Liberalization among Major World Trading Areas*. Cambridge, Mass.: MIT Press.
- Wolf, Martin. 2001. "Will the Nation-State Survive Globalization?" *Foreign Affairs* 80(1):178-90.
- World Bank. 2000. *Trade Blocs*. New York: Oxford University Press.
- . 2001. *Global Economic Prospects and the Developing Countries 2001*. Washington, D.C.: World Bank.
- . 2002a. *Globalization, Growth, and Poverty*. A Policy Research Report. New York: Oxford University Press.
- . 2002b. *World Development Report 2002*. New York: Oxford University Press.

Paths of Institutional Development: A View from Economic History

Karla Hoff

This article surveys an influential new research program on historical paths of institutional development and their consequences for growth. The research program exploits the experience of European colonialism as a kind of “natural experiment” whose results bear on the way institutions affect development. The central hypothesis of this research is that societies that began with more extreme inequality were more likely to develop institutions allowing much of the population only limited access to economic opportunities. The research has uncovered a striking reversal of fortune among the areas colonized by Europe; those that were relatively rich in the 1600s are today far poorer than the areas (such as the United States and Canada) that initially were viewed as relatively undesirable. The timing of the reversal—at the onset of the Industrial Revolution, when there was probably a premium on broad participation in commercial activity—suggests that institutions associated with high inequality may be a causal factor in low aggregate incomes. This research program is still in its early stages. But studies of institutions in India using data rich enough to permit hypothesis-testing provide evidence supporting the hypotheses developed in the analysis of the European colonial experience.

Economists have for some time recognized that institutions—the “rules of the game” that shape incentives and opportunities—are a key determinant of the wealth or poverty of nations.¹ But they are just beginning to understand the process of *institutional change*. In earlier work I surveyed recent developments that led to a consensus that the traditional neoclassical model, because it abstracts from institutional change, leaves out the heart of development economics (Hoff and Stiglitz 2001; Hoff 2001). This article surveys an influential new research program on historical paths of institutional development and their consequences for growth.

Much of the work discussed here is concerned with two puzzling facts about former European colonies (Engerman and Sokoloff 1997, 2002; Sokoloff and Engerman

2000; Acemoglu, Johnson, and Robinson 2002). First, many European colonies viewed by forecasters and migrants of the time as offering the best prospects for wealth are among the poorest countries in the world today. Second, the point in time at which these colonies fell behind was not during the early period of colonialism but at the onset of the Industrial Revolution. In simplest terms, the thesis put forward persuasively in recent work is that the factors that made these European colonies relatively wealthy in 1500, 1600, or 1700 also made possible a colonization strategy that created or perpetuated stark inequalities in wealth and political power—a strategy whose legacy is institutions that made these areas ill-suited for modern economic growth.

In 1700, Mexico and the colonies that were to become the United States had a very similar per capita income (based on approximate estimates), and the sugar-producing islands of Barbados and Cuba were far richer (table 1). In the 16th, 17th, and 18th centuries, the North American mainland was widely considered to offer relatively poor economic prospects compared with the vast opportunities in the Caribbean and Latin America. The colonial powers viewed Canada—famously characterized by Voltaire as “a few acres of snow”—to be of comparable value to the small sugar-producing island of Guadeloupe. But the United States and Canada ultimately proved to be far richer than other economies of the hemisphere. All Latin American countries in Table 1 fell behind the United States in per capita income in the 19th century, and Argentina, Mexico, and Peru fell still further behind in the 20th century.

In the modern period at least, few indicators are as strongly correlated with a country’s level of development as its level of urbanization (see, for example, Glaeser 1999). Moreover, longer time series are available for urbanization than for per capita

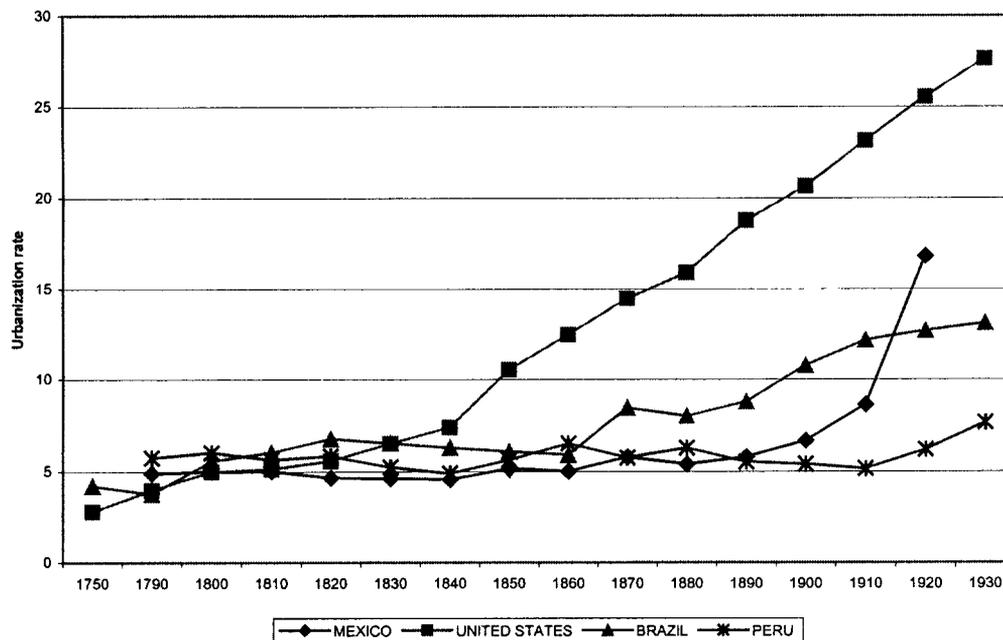
Table 1. Record of Per Capita Gross Domestic Product in Selected Economies of the Americas, Selected Years, 1700–1997

| <i>Economy</i> | <i>Per capita GDP as a percentage of U.S. value</i> | | | |
|--|---|-------------|-------------|-------------|
| | <i>1700</i> | <i>1800</i> | <i>1900</i> | <i>1997</i> |
| Argentina | — | 102 | 52 | 35 |
| Barbados | 150 | — | — | 51 |
| Brazil | — | 50 | 10 | 22 |
| Chile | — | 46 | 38 | 42 |
| Cuba | 167 | 112 | — | — |
| Mexico | 89 | 50 | 35 | 28 |
| Peru | — | 41 | 20 | 15 |
| Canada | — | — | 67 | 76 |
| U.S. per capita GDP (1985 U.S. dollars) | 550 | 807 | 3,859 | 20,230 |

Note: — Not available.

Source: Sokoloff and Engerman (2000).

Figure 1. Urbanization Rates in Four New World Economies, 1750–1930



Note: Urbanization rate is the percent of population living in urban areas with a population of at least 20,000.
Source: Acemoglu, Johnson, and Robinson (2001b). Reprinted with permission.

income. If urbanization is taken as a proxy for development, the U.S. economy diverged sharply from the rest of the hemisphere in the mid-19th century, just as the United States began to industrialize (figure 1). Per capita incomes throughout the Americas have increased greatly since 1900, but the gap between the United States and Canada, on the one hand, and Latin America, on the other, has changed little in proportional terms since 1900 (see table 1).

Why did the areas favored by the forecasters of the 1700s fall behind? Clearly development is not just about *having* productive opportunities. If that were the case, it would be hard to explain why all the Latin American colonies, which were generally richer in 1700 than the areas to the north, fell behind beginning in the 18th century. More than *having* productive opportunities, most economists would agree that development depends on being able to *create* a never-ending supply of new opportunities in the future. Economic historians Stanley Engerman and Kenneth Sokoloff (1997, 2002; Sokoloff and Engerman 2000) have argued that the reason that incomes in the United States and Canada diverged from those in the rest of the hemisphere was that a key to successful early industrialization was the *breadth of access to opportunities for social and economic advancement*—the ability to own land,

obtain schooling, borrow, and innovate. Among the colonies of the Americas, only the United States and Canada provided the social infrastructure—the collection of laws, institutions, and government policies (Hall and Jones 1997, 1999)—that made participation in investment and entrepreneurship possible for a broad segment of the population. At the end of the 18th century, the Industrial Revolution generated opportunities whose value depended on broad participation in entrepreneurship, investment, and innovation. Thus a “reversal of fortune” occurred as the United States and Canada surged ahead of societies in which a large fraction of the population was illiterate, disenfranchised, and without assets to borrow against.²

This article first gives an overview of the surprising systematic patterns in the historical paths of institutional development in New World economies and evidence that differences in the choices of the colonial powers relating to the institutions they set up have had a powerful effect on the performance of these economies in the past two centuries. Next it discusses two recent econometric studies that pinpoint the causes and consequences of inefficient institutions of particular types in rural India. Finally, it uses the example of patent institutions to illustrate the possibility that, for institutions, the devil is in the details—the fine details of an institution can make the difference between broad and narrow access to opportunity.

I. Paths of Institutional Development in the New World

The most systematic studies of paths of institutional development have focused on the Americas (see Engerman and Sokoloff 1997, 2002; Sokoloff and Engerman 2000; Acemoglu, Johnson, and Robinson 2002; and Easterly 1999, 2000). The patterns that this work discerns over the past 400 years point to two broad conclusions.

- European colonialism was a historical moment when factor endowments imposed constraints on the choices of a small number of colonial powers with respect to the institutions they set up. Factor endowments, broadly defined, were major influences on the level of economic and political inequality in the early days of colonial rule.
- The institutions established at the outset of colonial rule then influenced policies that shaped factor endowments (levels of physical and human capital) and institutional trajectories in ways that tended to reproduce the initial level of political and economic inequality. In colonies characterized by high initial inequality, the majority of people could not vote, own land, or obtain an education for hundreds of years. Nor, as Acemoglu, Johnson, and Robinson (2002) emphasize, could they enjoy effective protection from expropriation of their property.

Three Kinds of Factor Endowments

The central aspect of factor endowments that shaped initial inequality in the European colonies was labor scarcity.³ The extent of labor scarcity depended not only on the size of the population but also on its density (which influenced how easily the native population could be enslaved) and the local climate and soils (which influenced the profitability of importing slaves). Because there were no prohibitions against slavery in any of the regions colonized by the Europeans, areas where the climate and soils were suitable for plantation agriculture were those where a slave labor force was imported.⁴

Engerman and Sokoloff (1997) distinguish three broad classes of colonies by their factor endowments during the early period of colonization (1500–1650):

- *Dense populations of natives and rich supplies of minerals.* Many parts of Spanish America had large concentrations of populations of Native American descent that survived contact with the colonizers—and rich concentrations of minerals. Spain adopted a policy throughout Spanish America of distributing vast grants of land, including claims to the labor of the native population residing on the land. Spain also limited the immigration of European settlers, which contributed to the persistence of elites and the maintenance of vast landholdings even where production activities (as in Argentina) were not characterized by economies of scale.
- *Climate and soils well-suited to producing sugar and other highly valued crops cultivated with slave labor.* In the Caribbean, factor endowments were suitable for plantations based on slave labor, and scale economies supported the competitive success of large-scale sugar plantations. The reliance on slavery and the disparity in landownership made the distribution of wealth, income, and human capital extremely unequal, not only among the population as a whole but even among free men.
- *Dispersed indigenous population and climate and soils suited for grains and livestock.* In North America the preexisting population of Native American descent was very sparse. Except in the South, the colonies had soils and climates that made the production of crops using slave labor unprofitable. As a result, development depended on labor of European descent. Such labor was scarce because the North American colonies were not viewed as attractive places to settle and because grains could be profitably produced on very small farms (see Engerman and Sokoloff 2002:table 2).

Thus for different reasons the colonies in the first two categories were characterized almost from the outset by an abundance of slave or quasi-enslaved labor with very low human capital. Those in the third category were not. These factor endowments shaped the possible choices of the colonial powers with respect to the institutions

they set up. To concentrate power and wealth in the hands of a few and to extract labor and tribute from the rest proved possible and profitable in the first two classes of environments but not in the third.

The French in Canada and the British in North America began many settlements with an attempt to introduce a seigniorial system for landholdings. But because of the labor scarcity, the system either failed to take hold or collapsed after a short period (Galenson 1996). The economic historian David Galenson (1996:144) emphasizes the pivotal role of labor scarcity in the economic organization of the early colonies:

The extreme labor shortage...allowed many early settlers to gain their economic independence from the manorial lords, and establish separate farms. . . . Although the establishment of large estates to be worked by tenants and landless laborers was the initial model on which these proprietary colonies [Maryland, Jamestown, Carolina, New Jersey, and New York] were usually based, *the greater economic power conferred on settlers by the New World's labor scarcity prevented these English tenures and practices from effectively taking hold*, and proprietors were often forced to adapt by simply selling their land outright to settlers. (emphasis added)

In the Southern colonies of what became the United States, conditions were somewhat different. The Southern climate was suitable for crops—rice, cotton, and tobacco—that exhibited scale economies. But even here the slave population and the degree of inequality were much smaller than those in Latin America, in part because the South was unsuitable for sugar production and in part because institutions in the South were determined at the national level or by competition among states.

All the economies established in the Americas had abundant land relative to labor. But their climates, soils, minerals, and density of native populations differed in ways that meant that most regions were characterized almost from the outset of colonization by extreme inequality, whereas the colonies that were to become the United States and Canada were not.

Paths of Institutional Development

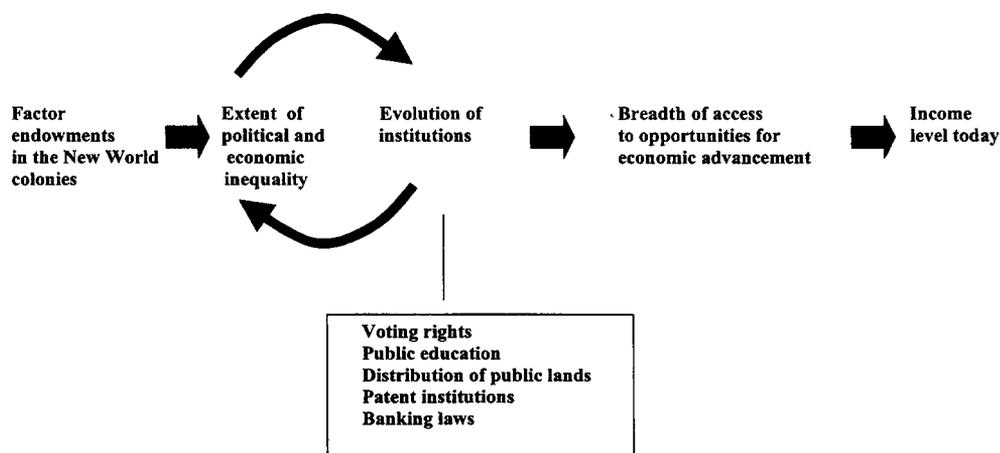
Why should inequality hundreds of years ago matter for per capita income today? What are the mechanisms that link the historical factor endowments to social organization and economic performance today? Recent and ongoing research has established systematic patterns in the institutional development paths of New World economies: In societies with high inequality at the outset of colonization, institutions tended to evolve in ways that restricted to a narrow elite access to political power and opportunities for economic advancement (Engerman and Sokoloff 1997, 2001, 2002; Engerman, Mariscal, and Sokoloff 1999).

Engerman and Sokoloff have discerned such a pattern in a wide range of public policies: the right to vote and to vote in secret, schooling, the distribution of public land and other natural resources, patent institutions, and banking laws. Moreover, they argue that high inequality at the outset of colonization through its effects on institutions may provide an important part of the explanation for the divergence in per capita income among New World economies. Their central thesis is worth quoting at some length (see figure 2 for a schematic representation of this thesis):

The factor endowment and the degree of inequality may influence the directions in which institutions evolve, but these institutions, in turn, can affect the evolution of the factor endowment and of the distributions of human capital, wealth, and political power. The initial conditions had long-lasting effects, however, not only because they were difficult to change, but also because government policies and other institutions tended generally to foster their persistence. (Engerman and Sokoloff 2002:63–64)

Voting Rights. One mechanism through which initial disparities in wealth and political power might be reproduced over time is limitations on suffrage and on the right to vote in secret. Most societies in the Americas were democracies by the middle of the 19th century, but they differed sharply in the breadth of effective access to the vote. Until early in the 19th century all countries, including the United States, limited the right to vote to white men with significant property. But by the mid-19th century, the share of the population voting in the United States and Canada was far greater

Figure 2. Institutional Origins and Persistence in New World Economies



Source: Author's illustration.

than that in Argentina, Brazil, Chile, and Ecuador (table 2). At the beginning of the 20th century the share voting in the United States and Canada was more than twice that in Argentina, one of the most progressive Latin American countries at the time.

In the United States, labor scarcity played a role in broadening voting rights. This is evidenced by the fact that the pioneers in extending the franchise were Western states competing for migrants (Engerman and Sokoloff 2001, 2002). Nearly all new entrants to the union in the 19th century extended voting rights to all white men.

Public education. Most Latin American countries did not provide public primary schooling on a scale sufficient to serve the general population until the 20th century. In contrast, by the early 19th century locally funded public primary schooling was widespread in the United States and Canada (although the southern United States,

Table 2. Voting and Literacy Rates in Selected Economies of the Americas, Various Years, 1850–1925

| <i>Period and country^a</i> | <i>Share of the population voting (%)</i> | <i>Literacy rate (%)^b</i> |
|---------------------------------------|---|--------------------------------------|
| <i>1850–1900</i> | | |
| Argentina, 1896, 1869 | 1.8 | 23.8 ^c |
| Brazil, 1894, 1872 | 2.2 | 15.8 ^d |
| Chile, 1869, 1865 | 1.6 | 18.0 ^d |
| Ecuador, 1856 | 0.1 | — |
| Canada, 1867, 1861 | 7.7 | 82.5 ^e |
| United States, 1850, 1870 | 12.9 | 80.0 |
| <i>1900–25</i> | | |
| Argentina, 1916, 1925 | 9.0 | 73.0 |
| Brazil, 1914, 1920 | 2.2 | 30.0 |
| Chile, 1920, 1925 | 4.4 | 66.0 |
| Colombia, 1918 | — | 32.0 ^f |
| Mexico, 1920, 1925 | 8.6 | 36.0 |
| Canada, 1911 | 18.1 | — |
| United States, 1900, 1910 | 18.4 | 92.3 |

Note: — Not available.

^aThe first year in each row relates to voting and the second to literacy. Where only one year appears, it relates to the indicator for which data are available.

^bExcept where otherwise specified, the literacy rate refers to the age group 10 and above.

^cAge 6 and above.

^dAge 7 and above.

^eAll.

^fAge 15 and above.

Source: Engerman and Sokoloff (2002:tables 7 and 8).

which had greater inequality and greater heterogeneity in the population, lagged behind). As a result, around 1870 the United States and Canada had literacy rates roughly four times those in many Latin American countries.

When literacy rates first diverged, the differences in per capita income between the United States and Canada and the rest of the hemisphere were small. So differences in aggregate resources to invest in schooling do not explain this pattern.

An alternative explanation focuses on political and economic inequality. If only the wealthy have the right to vote, all voters face tradeoffs. Providing mass education would raise the productive potential of the poor majority, which might promote growth and create positive spillovers for the wealthy. But the wealthy would bear a disproportionate share of the cost of that education, and the newly educated poor might agitate for the right to vote, which would threaten the political power of the elite. Among the New World economies, those with greater political inequality (as reflected in the share of the population voting) had a smaller share of the population enrolled in school, controlling for per capita income, time, and region (Engerman, Mariscal, and Sokoloff 1997). Among all economies today, those with greater income inequality tend to have lower spending on schooling, particularly primary schooling, and tend to be less democratic.⁵

Distribution of public lands. The governments of colonies or nations were regarded as the owners of the extensive public lands that existed into the 19th century and beyond. Public policies for transferring these lands to private hands were often the subject of political debates and struggles. The outcomes in the United States and Canada differed greatly from those in Latin America. The two North American countries awarded small landholdings to people who would settle and farm the land for a specified period. In contrast, Argentina and Mexico tended to award large landholdings to developers.

This policy difference led to extreme differences in the degree of inequality in rural landownership in these four countries at the beginning of the 20th century (table 3). In the two Latin American countries a small minority of households owned all the land—19 percent in Argentina and 2 percent in Mexico. In contrast, in Canada 87 percent of heads of household in rural areas owned land, and in the United States, 75 percent did. Throughout much of U.S. and Canadian history the vast majority of occupiers of farmland were landowners (Engerman and Sokoloff 2002).

Thus large differences existed between the United States and Canada, on the one hand, and the rest of the Americas, on the other, in public policies toward suffrage, education, and land. Based on these and similar patterns, Engerman and Sokoloff (1997, 2002) argue that initial differences in the degree of inequality—which can be attributed largely to factor endowments—had long-lasting effects on the paths of development of the economies of the Americas. In economies characterized by stark inequality at the outset of colonization, institutions evolved in ways that tended to

Table 3. Rural Landholding in Argentina, Mexico, Canada, and the United States, around 1900

| <i>Country and year</i> | <i>Share of the total population owning land (%)</i> |
|-------------------------|--|
| Argentina, 1895 | 19.2 ^a |
| Mexico, 1910 | 2.4 |
| Canada, 1901 | 87.1 |
| United States, 1900 | 74.5 |

Note: Landownership is defined as follows: in Argentina, the share of landowners in the male population ages 18–50; in Mexico, heads of household who own land; in Canada, occupiers of farmland who are owners; and in the United States, farms that are owner-operated.

^aThe average across the nine states of Argentina for which data are available.

Source: Engerman and Sokoloff 2002, table 6.

protect the elite and maintain a large class of poor, uneducated, and disenfranchised people. Neither the abolition of slavery nor the creation of democratic governments changed the basic patterns.⁶

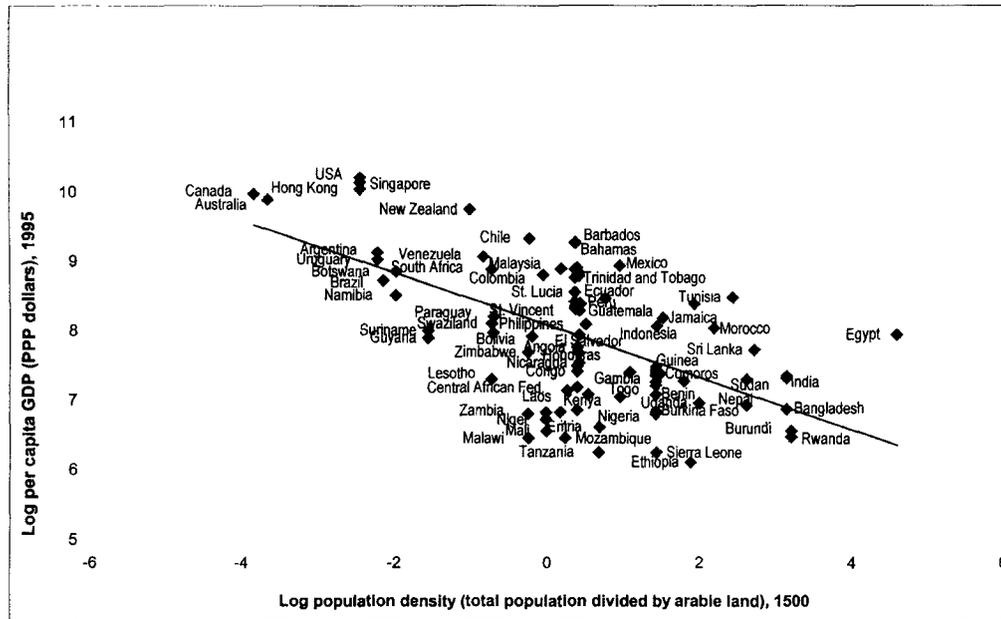
An Empirical Test

The work of Engerman and Sokoloff and Acemoglu, Johnson, and Robinson lays out a provocative theory: that the differences between the institutions that the European colonizers put in place (or took over) at the outset of colonization account for the reversal in relative incomes between 1500 or 1700 and the present day. A testable implication of the theory is that areas characterized by high labor density in 1500 should be relatively poor today. Figure 3 shows that this is indeed the case. There is a strongly negative relationship between population density in 1500 and per capita income today. The approximately linear relationship between these two variables, expressed in logarithms, means that there is a roughly constant relationship between a 1 percent increase in population density in 1500 and a 1 percent decline in per capita income today.

This statistical relationship is robust, remaining essentially the same for subsamples of colonies that are on the same continent, were colonized by the same country, are at the same distance from the equator, and exclude Australia, Canada, New Zealand, and the United States. Acemoglu, Johnson, and Robinson (2001b:2) offer the following explanation:

Relatively poor areas were sparsely settled, and this enabled or induced Europeans to develop settler colonies with institutions encouraging investment and commerce by a broad cross-section of the society. In contrast, a large population and relative prosperity made it profitable for the Europeans to

Figure 3. Population Density in 1500 and Per Capita Income in 1995 for Former European Colonies



Source: Acemoglu, Johnson, and Robinson (2002:figure 2). Reprinted with permission of MIT Press Journals.

set up extractive institutions, with political power concentrated in the hands of a small elite. High population density, for example, meant a large supply of labor that the Europeans could force to work in mines or plantations, or tax heavily by taking over existing tribute systems.

Acemoglu, Johnson, and Robinson (2002) go on to demonstrate that the economic growth of the initially sparsely settled colonies diverged sharply from that of the initially densely settled colonies at the onset of the Industrial Revolution (see figure 1, which illustrates the relationship for four economies of the Americas). This evidence—the “second fact” about the reversal of fortune—is consistent with the view that it was the interaction of national institutions and opportunities for industrialization during the 19th century that caused the divergence of incomes between the initially sparsely settled economies and those that were more densely populated at the outset of colonization.

The historical findings reported here are sharply at variance with two traditional views of economic development. One is the functionalist view that institutions are endogenous, flexible, and efficient. In this view, efficient institutions will “naturally”

emerge over time.⁷ Key institutions for the success of capitalist economies are secure property rights and the rule of law. But Acemoglu, Johnson, and Robinson (2002:table 7) find that colonies that started with dense populations were unlikely to develop effective property rights institutions. Among the former European colonies, areas with a high population density in 1500 not only are relatively poor today but also have relatively poor property rights institutions, as measured by the constraints on the executive and the risk of expropriation. The evidence supports the view that the stark political and economic inequality in some of the European colonies may have had a powerful effect on growth in the modern era through two distinct channels: direct policies that limited the breadth of access to social opportunity (see figure 2), and the insecurity of property rights engendered by the concentration of political and civil rights in a narrow elite.

The second view is that national heritage has a preeminent influence on the institutions of the former European colonies. In this view, the British colonies succeeded because of their heritage of the rule of law and constraints on the executive, and the colonies of Latin America fell behind because of their heritage of absolutist rule.⁸ One problem with this view is that the former British colonies include very poor countries—such as Belize, Guyana, and Jamaica—that resemble neighbors with similar factor endowments and no history of British rule. The theory described in this article suggests quite a different explanation for the comparative success of the former British colonies. The early colonizers of the New World—Spain and Portugal—chose the regions most heavily populated at the time, whereas Great Britain, as a late colonizer, was left with many of the less desirable lands, areas lacking dense populations or rich minerals. The need to solve the problem of labor supply in North America set the British colonies on a path of institutional development that diverged sharply from that followed in Latin America and the British sugar colonies. That path, rather than the national heritage, appears to explain the ultimate success of the North American colonies.

II. Tests of Causal Links among Inequality, Institutions, and Growth

The previous section set out the hypothesis that the level of inequality at the outset of colonization affected the evolution of institutions with strategic importance for growth. Given the limited data available for remote periods, subjecting this hypothesis to a test of causation is difficult. The richer data available for present-day developing economies permit stronger tests of causal links from the level of inequality to institutions and from institutions to growth. This section briefly describes the results of two tests using data from India.

Sugar Cooperatives in Maharashtra

Banerjee and others (2001) document and explain a puzzle in economic performance in the Indian state of Maharashtra over a recent 23-year period (1971–93). The study examines the performance of one present-day institution—sugar cooperatives. Cooperatively organized production of sugarcane has lower yields and lower growth in the fertile eastern region of the state than in the arid, less fertile western region. The surprising resolution of the puzzle illustrates the effects that the distribution of property rights can have on efficiency.

In the eastern region of the state, landholdings are heterogeneous in size. The wealthier members of the sugar cooperatives, who wield disproportionate control rights, set low prices for sugar supplied by members and divert the retained earnings to their own benefit. They prefer a lower price because it increases the rents they can extract. But a side effect of low sugar prices is weak incentives for farmers to improve productivity.

In the western region of the state, landholdings tend to be uniformly small. The cooperatives set the prices for sugar supplied by members near the world price. With high sugar prices, the farmers have strong incentives to improve productivity. Thus growth is higher and capacity more fully exploited than in the east.

Tenancy in West Bengal

Agricultural tenancy reform in the Indian state of West Bengal provides the setting for another test of the causal link between a present-day institution and efficiency. Before 1977, sharecropping contracts in West Bengal, involving around 2 million sharecroppers, generally assigned 50 percent of output to the tenant. In 1977, a new administration came into office with enforcement of a long-dormant tenancy law as one of its highest priorities. The law gave tenants a choice of registering with the government and stipulated that tenants who had registered could not be evicted from the land they sharecropped as long as they paid the landlord a minimum share-rent of 25 percent of output. Thus for most tenants, the reform increased their share of output on sharecropped land from 50 percent to 75 percent. It also gave them permanent, inheritable tenure on the land they sharecropped. In the decade after the reform, West Bengal achieved a breakthrough in agricultural productivity growth (table 4).

Banerjee, Gertler, and Ghatak (2002) use two approaches to measure the effect of the land reform on productivity. The first is a quasi-experimental approach that uses Bangladesh as a control. The second uses the fact that the land reform was implemented more intensively in some areas than in others to identify the effect of the change in tenants' share on agricultural production. The two approaches provide similar estimates, showing that the tenancy reform increased sharecropper yields by between 51 and 62 percent. These estimates imply that the tenancy reform explains

Table 4. Annual Growth in Food Grain Production in West Bengal and All of India, 1968–2001 (%)

| <i>Period</i> | <i>West Bengal</i> | <i>All of India</i> |
|---------------|--------------------|---------------------|
| 1968–81 | 0.43 | 1.94 |
| 1981–92 | 6.97 | 2.36 |
| 1992–2001 | 0.74 | 1.52 |

Source: For 1968–81, Banerjee, Gertler, and Ghatak (2002); for 1981–2001, Centre for Monitoring the Indian Economy (2002).

more than a quarter of the subsequent growth in agricultural productivity in West Bengal from 1979 through 1993. In that period, the rate of agricultural productivity growth in West Bengal rose from one of the lowest among Indian states to one of the highest. In the past decade, however, growth has slowed, for reasons not yet examined.⁹

III. Patent Institutions: Democratic and Undemocratic

As already described, an initial level of inequality tends to reproduce itself through such institutions as suffrage, schooling, land policies, and property rights institutions. But it is not just the broad policies of a government that influence access to economic opportunity. The fine details are also important. Rules that appear to be symmetric, because they apply equally to all, will not provide symmetric access if transaction costs affect different groups in the population differently. A comparison of the 19th-century patent systems in the United States and Great Britain provides an illustration. This comparison is used because of the richness of the data. But the kinds of costs and procedures that limited access to patent rights in 19th-century Britain were also important in Brazil and Mexico in the 19th century (Engerman and Sokoloff 1997; Khan and Sokoloff 2001).

The U.S. patent system of the 19th century was strongly influenced by British law. But the U.S. system made three innovations. First, it dramatically lowered fees. Then it gave inventors unrestricted freedom to assign their patent rights to others, which promoted a market for new technology and made it possible for individuals to become full-time inventors (Lamoreaux and Sokoloff 1999). Finally, it created impersonal administrative procedures for handling applications, which extended the protection of property rights to those with no special influence with government while reducing uncertainty about the value of inventive activity (table 5).

Not surprisingly, the differences in characteristics between the patent systems were associated with very different outcomes. The British system tended to restrict access to intellectual property rights to an elite—those with substantial wealth,

Table 5. Characteristics of Patent Institutions in the United States and Great Britain, 1790–1860

| <i>Characteristic</i> | <i>United States</i> | <i>Britain</i> |
|--|---|--|
| Patent fee | \$30–35 (20–30 percent of per capita income in 1790, and a declining share thereafter). Patent applications could be mailed to the Patent Office free of postage. | A multiple of per capita annual income, and 35–40 times the U.S. fee in real terms. |
| Freedom to assign rights to inventions | Unrestricted. | The number of assignees was limited and could be increased only by a private act of Parliament. |
| Application process | Impersonal, routine administrative procedure. | Processing by seven different offices was required. Many more offices were involved if patent protection was to extend to other British Isles besides England. The signature of the sovereign was required at two distinct stages. |
| Dispute settlement | Relatively predictable, since many parameters were established by statute. | Judges had jurisdiction in determining the usefulness of an invention and thus the validity of a patent. In doing so, judges exercised broad discretion. |
| Access of inventors to information on patents previously granted | All records were centralized in one office, to which access was free. | A fee was required to read a patent. Obtaining information was sufficiently difficult that patent agents were generally used as intermediaries. They were few in number and thus able to keep the prices for their services high. |

Source: Author's summary based on Sokoloff and Khan (1990) and Khan and Sokoloff (1998).

political connections, or technical knowledge. One measure of the breadth of access to intellectual property rights is the share of patentees who were merchants, professionals, and gentlemen. Among urban patentees in the United States, that share fell from 50 percent in the period 1790–1804 to 18.6 percent in 1836–50, whereas for all English patentees, this share remained roughly constant at 40 percent (table 6).

Another measure of the breadth of access to intellectual property rights is the share of patents granted to people with little previous record in invention. In the early 19th century, the share of patents granted to people who received a single patent over their career was 57.5 percent in the United States and 42.9 percent in Britain (table 6). In the United States, this share rose over the early 19th century, as “men with relatively common skills and knowledge [were pulled] into invention” (Sokoloff and Khan 1990:377). Indirect evidence suggests that such patentees were responsible for more than just low-value patents.

Table 6. Characteristics of Inventive Activity in the United States and the United Kingdom, 1790–1860

| <i>Characteristic</i> | <i>United States</i> | <i>Britain</i> |
|---|---|---|
| Share of patentees who were doctors, “gentlemen,” or merchants in the periods ^a : | | |
| 1790–1804 | 50.0 percent | 41.8 percent |
| 1805–1822 | 38.7 | 40.9 |
| 1823–1836 | 24.6 | 47.7 |
| 1836–1850 | 18.6 | 39.1 |
| Share of patents granted to individuals who received a single patent over their career, 1812–29 | 57.5 percent | 42.9 percent |
| Patents per capita, 1810–30 | 3.5 | 2 |
| Average education of patentees | Low. Even among those responsible for the most important inventions in 1790–1846, half had little or no formal schooling, and less than a quarter had attended college. | High |
| Composition of patent inventions | Balanced over sectors | Concentrated in capital-intensive sectors |

^aThe data for the United States are based on urban patentees only. Urban patentees received 31.3 percent of all patents in 1805–11 but only 22.1 percent in 1830–36 and 28.0 percent in 1836–42. The data therefore probably understate the difference between Britain and the United States in the share of patentees who are merchants, professionals, and gentlemen.

Source: Author’s summary based on Sokoloff and Khan (1990) and Khan and Sokoloff (1998).

In the United States, even the great inventors, those whom histories of technology credit with at least one important invention, were unexceptional in schooling and technical skills. For example, nearly half of the great inventors had little or no formal schooling, and less than a fourth attended college (table 6). For them as for the small inventors, the expansion of markets during early U.S. industrialization provided incentives for committing resources to inventive activity. The activity of the great inventors was influenced by the same market forces as those driving invention by ordinary patentees, contrary to the popular view of inventors as disinterested geniuses (Khan and Sokoloff 1993).

The broad access to intellectual property rights in the United States spurred an enormous increase in patenting activity in all sectors, in virtually all subregions, and across a broad spectrum of the population. Evidence also points to important interaction effects with public infrastructure. The building of canals, such as the Erie Canal, caused patenting activity in the vicinity to boom, as detailed maps of patent awards illustrate (Sokoloff 1988). This finding suggests a link between patenting and access to extensive markets, because waterways were the only low-cost means of long-haul transport until the 1830s. Sokoloff estimates that the extension of the inland waterways in the northeastern United States accounted for a 10–20 percent increase in U.S. patents per capita between 1805–11 and 1830–36. Rural counties especially realized large and rapid growth in patents per capita as they gained access to waterways. These findings suggest important synergies between extensive markets, effective property rights, and technological change.

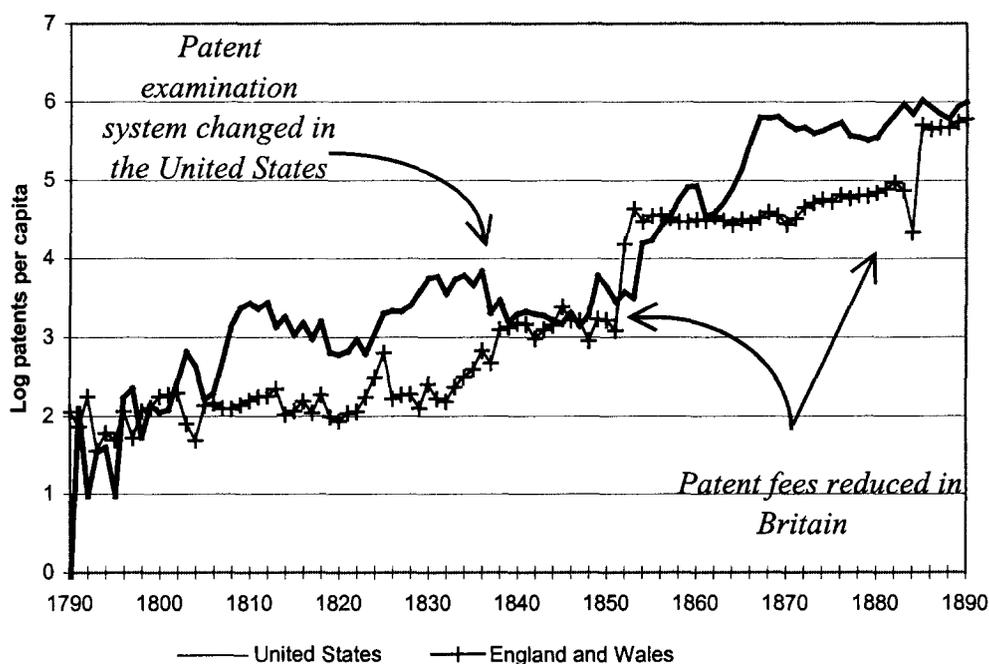
The comparison of U.S. and British patent institutions implies that the details of these institutions had an important influence on the breadth of effective access to intellectual property rights. In each country the same rules applied to everyone. But what looks symmetric in the British legal regime is not, because the system imposed needlessly high transaction costs, whose burden was much greater for the poor and the politically unconnected than for the rich.

The story has an epilogue. In 1852, one year after U.S. technology dazzled the world at the Crystal Palace exhibition in London, Britain made major changes in its patent system. In particular, it radically lowered patent fees. After these reforms the number of patents awarded in Britain jumped sharply, closing the gap with the United States in patents per capita (figure 4).

IV. Conclusion

European colonialism was a historical moment when factor endowments imposed important constraints on the choices of a small number of colonial powers with respect to the institutions they set up. This fact makes it possible to view that moment as an experiment that reveals the effects of differences in institutions. The experiment provides startling evidence of the impact of institutions on growth. It

Figure 4. Patents Per Capita in the United States and England and Wales, 1790–1890



Source: Khan and Sokoloff (1998). Reprinted with permission of Edward Elgar Publishing, Ltd.

shows that even in democracies, institutions can persist that exclude a broad segment of the population from opportunities to vote, to own land, to obtain schooling, and to secure effective protection of property rights. The reversal of fortune—when the economies that had been relatively poor surged ahead of those that had been relatively rich—occurred at the onset of the Industrial Revolution, when there was probably a premium on broad participation in entrepreneurship and innovation.

As Engerman and Sokoloff (2002:46) underscore, a “breathtaking” counterfactual is implicit in treating the moment of European colonialism as a natural experiment. In this counterfactual, but for the accident of high population density in 1500 or soils and climate suitable for sugar cultivation, areas of the Americas that are poor today would tend to be rich. Economists do not yet know what the key mechanisms of institutional change are, and they undoubtedly depend on the context. But viewing the early colonial period as a natural experiment suggests two likely drivers of institutional evolution: the level of political inequality and the scarcity of labor. These factors influence policies and institutions that in both their broad aspects (such as whether communities support public schooling) and their fine details (such as those of the British and U.S. patent systems) affect the breadth of access to opportunities for social advancement.

Although the work described here points to powerful and persistent forces underlying institutions, this work should not be taken to imply that the past is destiny. Institutions are malleable; factor endowments are malleable. Both depend on many influences. When a world exhibition of inventions provided striking evidence that Britain was falling behind in invention, it democratized its patent system. When the balance of political power shifted in favor of disadvantaged groups in West Bengal, India, in 1977, the state enforced a long-dormant tenancy law that expanded the economic rights of tenants and led to a big increase in productivity.

The nature of political competition and the balance between central and local control over government also influence institutions. In much of Latin America, national rather than local initiative finally expanded access to primary education (Engerman and Sokoloff 2002). In the United States, national forces determined many of the institutions of the South. In northeastern Brazil in recent years, a series of small reforms by a state governor created a new kind of competition among municipal governments that led to a radical change in the incentives of municipal health workers and major success in delivering health services to the poor (Tendler 1997:ch. 2).

The central questions in the research program described here are these: What are the forces that sustain inefficient institutions? How large are the effects of changing certain institutions? When can a change in institutions be sustained? Knowing the answers to these questions would take us a long way toward understanding both institutional change and the interventions that promote it. The case studies of rural institutions in India provide an example of the kind of detailed econometric work that will help advance understanding of the scope for interventions to change institutions. The research program suggests that the place to look for the obstacles to economic growth are institutions that protect narrow elites at the expense of broad access to opportunities for economic advancement and effective protection of property rights.

Notes

Karla Hoff is a senior research economist in the Development Research Group at the World Bank. Her e-mail address is: khoff@worldbank.org. The author is indebted to Gunnar Eskeland and Kenneth Sokoloff for helpful discussions, and to Zmarak Shalizi and the editors of this journal for the invitation to write this paper. She thanks Simon Johnson for assistance with Figures 1 and 3.

1. See, for example, North (1981, 1990) and Hall and Jones (1997). See World Bank (2001, 2002) for references to the vast literature.

2. The phrase *reversal of fortune* is from Acemoglu, Johnson, and Robinson (2002). Using urbanization and population density as proxies for income in 1500, the authors show that the reversal of fortune holds generally for New World colonies and for all former European colonies and that it does not hold for Europe and for all noncolonies. A parallel to the view in the 1700s of Canada as a few acres of snow—one whose prospects were quite dim—is the description of Australia as “an unchanging Paleolithic backwater” (McEvedy and Jones 1978:322, as cited in Acemoglu, Johnson, and Robinson 2002:1256).

3. This follows Engerman and Sokoloff (1997, 2002) and Acemoglu, Johnson, and Robinson (2002). Acemoglu, Johnson, and Robinson (2001a) put forward an alternative thesis: the key aspect of factor endowments was the disease environment. They argue that where prospective European settlers faced a high risk of disease, settler communities did not form and so the colonial powers did not wish to create a rule of law (broad access to secure property rights). Many scholars have raised concerns about this provocative thesis. One concern is that it fails to take into account that factor endowments, broadly defined, also imposed constraints on the ability of the colonial powers to set up a highly exploitative form of social organization: If slavery was unprofitable and if almost any settler could make a good living by farming on his own, highly inegalitarian institutions could not take root. In addition, Engerman and Sokoloff (2002) have argued that the pattern of European migration to the colonies does not support the hypothesis that the colonies with high mortality were unattractive to settlers. Finally, that areas with better institutions tend to have lower mortality makes a causal interpretation from 19th-century mortality figures to institutions problematic.

4. Slavery is just one example of an institution that influences labor scarcity. Another is the caste system in India, which barred all members of the lowest castes (shudras and “Untouchables”) from owning property and thus becoming self-employed cultivators. Many other examples are provided by Binswanger, Deininger, and Feder (1995).

5. Benabou (2000) provides a theoretical model of the tradeoffs that voters face. Easterly (1999, 2001) tests the implications empirically. Acemoglu and Robinson (2000) present a model in which political constraints block the adoption of policies that would increase wealth. For an accessible overview of the ways in which social polarization can hurt growth, see Easterly (2001:ch. 13).

6. North (1990) has suggested that the past matters because it shapes not only present opportunities but also ideologies and expectations. A characteristic expression of stark political inequality is an ideology that stigmatizes the subordinate group. In a controlled experiment, Hoff and Pandey (2003) find that the performance of low castes when they undertake a task for money is debilitated by the public announcement of caste and that the influence of caste may be mediated by distrust that effort will be rewarded.

7. Douglass North adopted this position in his early work (North and Thomas 1970) but abandoned it in his later work (North 1990). Broadly speaking, this view underlies the traditional neoclassical model, which assumes that fundamental forces of endowments, preferences, and technology drive allocations, history does not matter, and economic outcomes are the same as those that would emerge as the equilibrium allocation under a competitive market system.

8. A parallel explanation focuses on religious heritage. In one version, the distinctiveness of the North American colonies was due to their Protestant heritage. A remarkable example that again points to the importance of factor endowments and the much lesser importance of religion is explored in Kupperman’s (1993) comparative study of two Puritan colonies, the Massachusetts Bay Colony and Providence Island off the coast of Nicaragua.

9. These two tests provide evidence of a causal link between inequality and institutions and between institutions and growth but do not bear on the third aspect of the work by Engerman and Sokoloff and Acemoglu, Johnson, and Robinson—institutional overhang, the persistent effect of institutions that have been formally changed. Banerjee and Iyer (2002) test for institutional overhang in their work on agricultural productivity differences across districts in India. In some areas, the British colonial regime conferred *de facto* governmental powers on a landlord by giving him the right to collect taxes and retain a substantial share of the revenue. In other districts, landlords were bypassed in favor of village councils or direct taxation of cultivators, and the cultivators were given implicit property rights to the land. The land revenue system introduced by the British ended in 1947—independent India—does not impose direct taxes on agricultural incomes. Yet Banerjee and Iyer find that districts that were controlled by landlords have lower current agricultural productivity today, stemming from lower rates of investment and lower use of modern inputs, than do areas where property rights to the land were given to the cultivators.

References

- Acemoglu, Daron and James A. Robinson. 2000. "Political Losers as a Barrier to Economic Development." *American Economic Review Papers and Proceedings* 90(2):126–30.
- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001a. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91(5):1369–401.
- . 2001b. *Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution*. NBER Working Paper 8460. National Bureau of Economic Research, Cambridge, Mass.
- . 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution." *Quarterly Journal of Economics* 117(4):1231–94.
- Banerjee, Abhijit and Lakshmi Iyer. 2002. "History, Institutions and Economic Performance: The Legacy of Colonial Land Tenure Systems in India." Massachusetts Institute of Technology, Department of Economics, Cambridge, Mass.
- Banerjee, Abhijit, Paul Gertler, and Maitreesh Ghatak. 2002. "Empowerment and Efficiency: Tenancy Reform in West Bengal." *Journal of Political Economy* 110(2):239–80.
- Banerjee, Abhijit, Dilip Mookherjee, Kaivan Munshi, and Debraj Ray. 2001. "Inequality, Control Rights, and Rent Seeking: Sugar Cooperatives in Maharashtra." *Journal of Political Economy* 109(1):138–90.
- Benabou, Roland. 2000. "Unequal Societies: Income Distribution and the Social Contract." *American Economic Review* 90(1):96–129.
- Binswanger, Hans, Klaus Deininger, and Gershon Feder. 1995. "Power, Distortions, Revolt and Reform in Agricultural Land Relations." In T. N. Srinivasan and Jere Behrman, eds., *Handbook of Development Economics*, Vol. 3B. Amsterdam: Elsevier Science.
- Centre for Monitoring the Indian Economy. 2002. "Agriculture." New Delhi, India.
- Easterly, William. 1999. "The Middle Class Consensus and Economic Development." Policy Research Working Paper 2346. World Bank, Development Research Group, Washington, D.C.
- . 2000. "It's Not Factor Accumulation: Stylized Facts and Growth Models." World Bank, Development Research Group, Washington, D.C. Online document available at www.worldbank.org/html/prdmg/grthweb/stylized.htm.
- . 2001. *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics*. Cambridge, Mass.: MIT Press.
- Engerman, Stanley L. and Kenneth L. Sokoloff. 1997. "Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies: A View from Economic Historians of the United States." In Stephen Haber, ed., *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800–1914*. Stanford, Calif.: Stanford University Press.
- . 2001. *The Evolution of Suffrage Institutions in the New World*. NBER Working Paper 8512. National Bureau of Economic Research, Cambridge, Mass.
- . 2002. "Factor Endowments, Inequality, and Paths of Development among New World Economies." *Economia* 3(1):41–109.
- Engerman, Stanley L., Elisa V. Mariscal, and Kenneth L. Sokoloff. 1997. "The Persistence of Inequality in the Americas: Schooling and Suffrage, 1800–1945." University of California at Los Angeles, Department of Economics.
- Galenson, David W. 1996. "The Settlement and Growth of the Colonies: Population, Labor and Economic Development." In Stanley L. Engerman and Robert Gallman, eds., *The Cambridge Economic History of the United States*, Vol. 1. New York: Cambridge University Press.
- Glaeser, Edward L. 1999. "The Future of Urban Research: Nonmarket Interactions." Harvard University, Department of Economics, Cambridge, Mass.

- Hall, Robert E., and Charles I. Jones. 1997. "Levels of Economic Activity across Countries." *American Economic Review* 87(May):173–77.
- . 1999. "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114:83–116.
- Hoff, Karla. 2001. "Beyond Rosenstein-Rodan: The Modern Theory of Coordination Problems in Development." In Boris Pleskovic and Nicholas Stern, eds., *Annual World Bank Conference on Development Economics 2000*. Washington, D.C.: World Bank.
- Hoff, Karla and Priyanka Pandey. 2003. "Why Are Social Inequalities So Durable? An Experimental Test of the Effect of Indian Caste on Performance." World Bank, Development Research Group, Washington, D.C.: and Pennsylvania State University, Department of Economics, College Park, Pa.
- Hoff, Karla and Joseph E. Stiglitz. 2001. "Modern Economic Theory and Development." In Gerald Meier and Joseph E. Stiglitz, eds., *Frontiers of Development Economics*. Oxford: Oxford University Press.
- Khan, B. Zorina and Kenneth L. Sokoloff. 1993. "'Schemes of Practical Utility': Entrepreneurship and Innovation among 'Great Inventors' in the United States, 1790–1865." *Journal of Economic History* 53(2):289–307.
- . 1998. "Patent Institutions, Industrial Organization and Early Technological Change: Britain and the United States, 1790–1850." In Maxine Berg and Kristine Bruland, eds., *Technological Revolutions in Europe*. Cheltenham, U.K.: Elgar.
- . 2001. "The Innovation of Patent Systems in the Nineteenth Century: A Historical Perspective." University of California at Los Angeles, Department of Economics.
- Kupperman, Karen Ordahl. 1993. *Providence Island, 1630–1641: The Other Puritan Colony*. New York: Cambridge University Press.
- Lamoreaux, Naomi R. and Kenneth L. Sokoloff. 1999. *Inventive Activity and the Market for Technology in the United States, 1840–1920*. NBER Working Paper 7101. National Bureau of Economic Research, Cambridge, Mass.
- McEvedy, Colin and Richard Jones. 1978. *Atlas of World Population History*. New York: Facts on File.
- North, Douglass C. 1981. *Structure and Change in Economic History*. New York: Norton.
- . 1990. *Institutions, Institutional Change and Economic Performance*. New York: Cambridge University Press.
- North, Douglass C. and Robert Paul Thomas. 1970. "An Economic Theory of the Growth of the Western World." *Economic History Review* 2d ser., 23:1–17.
- Sokoloff, Kenneth L. 1988. "Inventive Activity in Early Industrial America: Evidence from Patent Records, 1790–1846." *Journal of Economic History* 48(4):813–50.
- Sokoloff, Kenneth L. and Stanley L. Engerman. 2000. "History Lessons: Institutions, Factor Endowments, and Paths of Development in the New World." *Journal of Economic Perspectives* 14(3):217–32.
- Sokoloff, Kenneth L. and B. Zorina Khan. 1990. "The Democratization of Invention during Early Industrialization: Evidence from the United States, 1790–1846." *Journal of Economic History* 50(2):363–78.
- Tendler, Judith. 1997. *Good Government in the Tropics*. Baltimore, Md.: Johns Hopkins University Press.
- World Bank. 2001. *World Development Report 2000/2001: Attacking Poverty*. New York: Oxford University Press.
- . 2002. *World Development Report 2002: Building Institutions for Markets*. New York: Oxford University Press.

IMF Conditionality and Country Ownership of Adjustment Programs

Mohsin S. Khan • Sunil Sharma

This article uses finance and agency theory to establish two key propositions about International Monetary Fund (IMF) conditionality and country ownership of IMF-supported adjustment programs. First, the authors propose that the conditionality attached to these programs is justified. Second, the article hypothesizes that country ownership of these programs is crucial for their success. Because IMF conditionality and country ownership are both necessary, the challenge is designing conditionality that maximizes ownership while providing adequate safeguards for IMF lending. The article analyzes several recent proposals aimed at enhancing country ownership of policies contained in IMF-supported programs. These proposals include encouraging countries to design their own adjustment and reform programs, streamlining structural conditionality, introducing flexibility in the timing of structural policy measures (floating tranche conditionality), and applying conditionality to outcomes rather than policies (outcomes-based conditionality).

International Monetary Fund (IMF) lending in support of adjustment programs is conditional on the borrowing countries adopting certain agreed policies. The conditions attached to these loans are commonly referred to as *IMF conditionality*. As the literature on the subject shows, discussions of the nature and merits of IMF conditionality have a long history (see, for example, Williamson 1983; Polak 1991; Guitián 1995; James 1996; and Boughton 2001:ch. 13). The issue has recently gained renewed attention, with questions raised about whether the conditions the IMF imposes on borrowing countries have been too intrusive and whether the design and implementation of IMF conditionality have undermined country ownership of adjustment programs aimed at correcting macroeconomic imbalances.

This viewpoint has gained considerable currency as a result of the capital account crises in Mexico, East Asia, the Russian Federation, and Brazil in the 1990s. But during such crises, which require rapid responses, there may not be enough time to secure full

country support for all the policy actions needed. Moreover, in the middle of a serious currency and financial crisis there may be greater agreement between the country authorities and the IMF on the immediate problems facing the country and the short-run measures needed to address them. Conditionality and ownership issues are probably less relevant for the recent capital account crisis cases and more relevant in standard IMF programs dealing with current account crises in low-income countries—and that is where most attention has been focused. The IMF itself has engaged in a comprehensive analysis of conditionality and ownership issues (see IMF 2001a.)

This article draws on finance and agency theory to establish two basic propositions about IMF conditionality and country ownership of adjustment programs. First, some form of conditionality exists in all borrower–lender relationships: Key to the ability to borrow is the ability to pledge income back. The IMF must have assurances that it will be repaid, and this requires that it place conditions on its loans. The analysis here is designed to dispel the fairly widespread notion—articulated, for example, by Diaz Alejandro (1984)—that conditionality stems from a “patron–beneficiary” relationship between the IMF and borrowing countries. As he puts it: “This is the key justification for ‘conditionality’: if you ask for a gift, you must listen to your patron” (Diaz Alejandro 1984:7). Finance considerations alone provide the justification for conditionality being a necessary part of IMF lending. Thus the view expressed by Killick (1997) that IMF conditionality should be the exception rather than the rule is incorrect. Indeed, this article argues that it should be exactly the reverse.

Second, country ownership of programs is essential, because it aligns the incentives of the borrower and the lender. For the borrowing country, program ownership is critical because without a firm commitment from the government and other relevant constituencies, the difficult policy measures needed to correct economic problems are less likely to be implemented. For the IMF, country ownership increases the probability that programs will succeed and so augments the protection of its resources provided by conditionality. Thus the Poverty Reduction Strategy Papers (PRSPs) that have been developed jointly by the IMF and the World Bank put substantial emphasis on country ownership. So, both IMF conditionality and country ownership have a clear rationale; the challenge is reconciling the two. This article considers various recent proposals for achieving such a reconciliation.

I. IMF Lending and Conditionality

This section examines the main features of IMF lending by comparing them to private loan contracts. It then assesses the implementation and effectiveness of IMF conditionality.

Conditionality in Private Financial Contracts

Between every borrower and lender there is a fundamental asymmetry in information. Borrowers always know more about their abilities, opportunities, and intentions than do lenders. This information asymmetry gives rise to two incentive problems: adverse selection and moral hazard.

Adverse selection arises before a transaction occurs and stems from the fact that information deficiencies make it difficult for lenders to distinguish good risks from bad. The IMF faces a different selection problem than private lenders in that only members experiencing distress approach it for financing and all have a right to its resources. Moral hazard arises after a lender has given funds to a borrower. Having obtained the funds, it may be in the borrower's interest to take risks that may raise returns—but also increase the likelihood of default.¹ In the financial world, contract designs and collateral, transparency, and reporting requirements attempt to mitigate such moral hazard. Monitoring by shareholders, debt holders, market analysts, rating agencies, and independent company directors serves the same purpose. Such monitoring is costly to firms, but it is done to ensure investors that their claims will be respected (Greenbaum and Thakor 1995; Mishkin 1998).

Pledging collateral is expensive, but firms incur that cost to provide lenders with assurance that borrowed funds will be used for stated purposes and in ways that will not jeopardize their eventual repayment. Indeed, the very existence of financial markets depends on such assurance. To that end, corporate governance and institutions and practices of finance share the common rationale of increasing resources that can be pledged to outside debt and equity holders while maintaining appropriate incentives for managers and workers. Collateral is provided so that in the event of default or if the borrower does not live up to the terms of the contract, the lender can recover its resources by taking control of the pledged assets and selling them.

In new or emerging firms with little physical or financial collateral, relinquishing control rights to venture capitalists provides the needed assurance that funds will be well spent, because the venture capitalists have considerable say in decisionmaking. The allocation of control rights to investors should be designed to provide maximum assurance to investors without impairing the functioning of the firm and its ability to exploit commercial opportunities. Even so, in many circumstances and for many firms—especially those in distress or those that have little reputational or financial capital—relinquishing such control rights may be costly and will almost certainly impose substantial limits on management discretion.

Generally, the imposition of such conditionality—the allocation of return streams, liquidity, and control rights—is made contingent on the evolution of the borrowing firm's balance sheet and follows a simple carrot-and-stick logic. Posting collateral that can be sold by the lender in case of default provides the borrower with an incentive to prevent default. Agency considerations dictate that the transfer of

collateral rights to the lender be made contingent on default or on observable measures of financial and nonfinancial performance.

Such conditionality also serves another purpose when it comes to control rights. After an adverse shock, a borrowing firm is more likely to gamble for resurrection—that is, engage in riskier behavior—the greater is the deterioration of its financial health. Contingent transfer of control rights protects investors in that it prevents firms from undertaking excessively risky activities to repair balance sheets. Hence it serves two purposes. Ex ante, it provides incentives for preventing default. Ex post, it constrains firms' ability to take gambles.

Even if a firm does not engage in risky strategies when its performance deteriorates, the contingent transfer of control rights provides the lender with the option of reexamining the new situation. For example, in a new enterprise the venture capitalist obtains full control if the firm performs poorly—whereas if the firm is profitable, the venture capitalist may retain cash-flow rights but agree to cede control and liquidation rights to the entrepreneur (Kaplan and Stromberg 2000).

IMF Conditionality

IMF lending and its associated conditionality follow broadly the same principles as private financial contracts, though several additional dimensions make IMF lending qualitatively different. The IMF is mandated by its Articles of Agreement to extend temporary financial assistance to member countries facing balance of payments difficulties “under adequate safeguards” (Article I). Like any lender, the IMF thus needs assurances from its borrowers that the funds lent to them will be used for the purposes defined by the Articles of Agreement and in a manner that does not jeopardize their contractual servicing and repayment. Many of the finance propositions relevant to private financial institutions also apply to the IMF.

A key aspect of IMF lending is that countries in need of IMF loans generally do not possess internationally valuable collateral. If they did, they could use it to borrow from private lenders and would not require IMF resources. In a crisis situation, however, even if a country had internationally acceptable collateral, it still might not be able to use it to borrow from private capital markets. There is an important difference between national income and income that can be pledged to foreign lenders, including the IMF. Foreign loans can be used to produce both tradable and nontradable outputs, but foreigners typically have no demand for a country's nontradables. In the absence of collateral, private loan contracts typically would include various covenants coupled with monitoring. Formally, a covenant is designed to protect the lender and prohibit the borrower from taking actions that could reduce the probability of repayment. Covenants can impose clear obligations on the borrower, impose limitations on or prohibit certain actions, and specify when a borrower is considered to be in default.

IMF conditionality can be viewed as a complex covenant written into the loan agreement. The policy prescriptions in IMF-supported programs essentially provide safeguards that the country will be able to rectify its macroeconomic and structural imbalances and will be in a position to service and repay the loan.² Thus the conditionality associated with IMF programs can be viewed as a substitute for collateral.

Conditionality attached to sovereign lending has a long history. For example, in 1818 Prussia, effectively bankrupted by the Napoleonic wars, approached Nathan Rothschild (of the House of Rothschild) for a loan and was asked to pledge “Prussian royal domains” as collateral (Ferguson 1998).

For many analysts the modern model for conditional lending to sovereign governments in the absence of collateral is the Turkish agreement of 1881—known as the Decree of Mouharrem—which was implemented after the Turkish government defaulted on its foreign debt in 1875. The decree created the Council of Ottoman Debt, comprising seven members that represented different groups of bondholders. A large portion of the Turkish government’s revenue was placed under the council’s direct control and used by it to service and repay the debt (Anderson 1966).

The League of Nations also attached strict conditions to its adjustment programs (or “reconstruction schemes,” as they were known) for six European countries in the 1920s. These conditions included maintenance of fiscal equilibrium and monetary discipline as well as currency reform. The League developed various means to enforce the programs and safeguard the interests of foreign creditors and bond holders, including appointing a commissioner to each country and an adviser to each central bank to administer and monitor the programs (Santaella 1993; James 2001).

Agreeing to IMF conditionality is an imposition on a country even though the two may share the same objectives—external viability, price stability, sustained high growth, reduced systemic risk, and so on. To be sure, the covenant in the loan agreement required by the IMF is far more complex and has different characteristics than covenants in simpler private financial transactions. It may also not always be part of the explicit contract, but it is always part of the implicit one, regardless of the language used in the documentation. The challenge in designing IMF conditionality is to specify the optimal covenant—the best policy conditions given the circumstances of the country, the disbursement intervals for the loan, the type of monitoring involved, and so on—to achieve program objectives while providing sufficient safeguards to the IMF. Even so, at a general level, IMF lending conforms to the principles governing private lending.

At the same time, IMF lending differs in significant ways. First, as noted, defining conditionality in IMF lending is much more complicated than in private financial transactions, where covenants may be quite straightforward. It basically involves assessing the macroeconomic imbalances or structural deficiencies that led to a country’s macroeconomic problems, then negotiating an agreement with country authorities that will address those issues.

Second, it is difficult if not impossible to establish the value of IMF conditionality. The value of negotiated conditionality largely depends on the degree to which the authorities of the borrowing country adopt the program and are willing to expend effort and political credibility to implement it.

Third, unlike private lenders—for whom it may be sufficient to deal with a firm's management—the IMF faces what in agency theory is called “moral hazard in teams” (Holmström 1982). This refers to situations in which a principal's payoff depends on the joint efforts of two (or more) agents. Typically, the detailed negotiations for an IMF-supported program are conducted with certain government representatives (central bank, finance ministry), whereas the success of the program depends on many other stakeholders in society—other ministries, trade unions, professional associations, civic groups, and nongovernmental organizations.

Fourth, the IMF is by design a cooperative that makes loans to its sovereign members. In the event of default there is no court to which it can appeal and no tangible collateral that can be used to cover lost resources. The enforcement mechanism for ensuring that borrowing countries live up to their obligations essentially amounts to a combination of moral suasion, maintenance of the borrower's reputation, peer pressure, and the threat of being shut out of international capital markets. Unlike private firms—where lending is generally subject to well-defined legal codes that can be enforced in courts and where shareholders or investors can change the management of a company or take it over—the IMF obviously cannot replace sovereign governments and can only refuse to provide financing to a country in arrears.

Finally, relative to private lenders the IMF, given its mandate and cooperative structure, faces what in other contexts has been called the “Samaritan's dilemma” (Buchanan 1975; Lindbeck and Weibull 1988). For example, to provide the right incentives, *ex ante* a private lender may impose harsh conditions on a borrower in the event of poor performance, but *ex post* (that is, after poor performance) it may not want to impose those conditions for a number of reasons. For instance, selling certain assets (or liquidating the company) may be worth less than keeping the assets with the company (or maintaining the company as a going concern).

But the IMF faces a different incentive problem because a borrowing country is always more valuable as a going concern. Hence country authorities know that in the event of poor performance, at worst the program will be renegotiated (Drazen and Fischer 1997). This creates the wrong incentives, *ex ante*. Countries know that faced with poor performance and a weak economy, the IMF is unlikely to impose harsh conditionality *ex post* because it has to be concerned about the welfare of borrowing countries. Being a cooperative institution, the IMF cannot simply walk away and cut its losses. Thus *ex ante* penalties have limited credibility because they are unlikely to be enforced *ex post*. This is one reason it has been suggested that the IMF should lend only to prequalified countries with good policy environments. In other words, the IMF should lend only to countries that have

good economic track records and that can provide good collateral (see Meltzer and others 2000).

Implementation of IMF Conditionality

Conditionality is implemented through program design and monitoring that tracks whether agreed policies are implemented in a timely and effective manner. Program design begins with in-depth analysis of the sources of a country's macroeconomic imbalances. (See Mussa and Savastano 1999 for a recent description of the IMF's approach to economic stabilization.) The next step is to agree with the authorities on policy objectives and on macroeconomic policies and structural reforms to achieve those objectives.

Monitoring takes various forms, depending on the borrowing country's circumstances and the IMF lending facility used. It generally includes prior actions, performance criteria, macroeconomic and structural benchmarks, and reviews. Monitoring also prohibits actions inconsistent with the IMF's Articles of Agreement, such as the introduction of new foreign exchange restrictions. Releases of IMF financing are linked to compliance with monitoring arrangements.

Prior actions are required when up-front implementation is critical to program success or to allay doubts about the authorities' commitment. Examples of required actions include passing an agreed budget, realigning the exchange rate, adopting structural reforms, or enacting relevant laws. Such actions must be taken prior to the approval of a program by the IMF's Executive Board, which then triggers the first disbursement.

Performance criteria normally include quantitative targets for specified financial aggregates (such as bank credit, net international reserves, or fiscal balance) and often include structural measures (such as tariff reductions, tax system revisions, or privatization of public enterprises). Meeting these criteria triggers the release of subsequent tranches of committed IMF resources.

Macroeconomic and structural benchmarks set targets for macroeconomic variables and structural policies important for effective program implementation. They do not directly affect scheduled disbursements.

Finally, reviews are used to assess overall progress toward program objectives, identify any sources of slippage (resulting from lack of policy implementation, external shocks, or program design issues), and take corrective actions. Reviews are usually specified as performance criteria—precluding further disbursements if the review is not concluded by the scheduled review date.

A common question is whether IMF conditionality works. On the one hand, Meltzer and others (2000:35) argue that “numerous studies on the effects of IMF lending have failed to find any significant link between IMF involvement and increases in growth or income.” But the IMF, not surprisingly, disagrees. The IMF's mandate is to

provide short-term lending to support balance of payments adjustments, and it believes that conditional lending has generally improved the external accounts of borrowing countries.

What is the evidence on the effectiveness of IMF conditionality? Specifically, have IMF-supported adjustment programs achieved their objectives of improving current account balances, increasing international reserves, lowering inflation, and raising growth? This is essentially an empirical question that requires evaluating the effects that past programs have had on the macroeconomic variables of interest. Such evaluations are conducted periodically by the IMF's Policy Development and Review Department and by its recently created Independent Evaluation Office, with the results reported to the IMF's Executive Board. In addition, a number of studies over the past 20 years, both inside and outside the IMF, have examined the question using a variety of empirical methods.

Almost all the empirical studies surveyed by Haque and Khan (2002) show that IMF-supported programs have improved current account balances and the overall balance of payments in borrowing countries. The results for inflation are less clear. Most of the studies indicate that although inflation usually falls, the decline is generally statistically insignificant. For growth, output is generally depressed in the short run—that is, during the stabilization phase—but increases as macroeconomic stability is established. Overall, the most recent empirical results in particular indicate that on average, IMF-supported programs and the conditionality they incorporate have been reasonably effective in achieving their main macroeconomic objectives.

II. Country Ownership of Programs

The case for country ownership of IMF-supported adjustment programs has a strong theoretical foundation. In the context of agency theory, principal-agent problems arise in situations in which one party (the principal) relies on the other (the agent) to achieve certain objectives. Due to information asymmetries and lack of perfect monitoring technology, if the agent's actions and their results cannot be easily verified and monitored, the agent has greater scope for pursuing its own interests rather than those of the principal.

But principal-agent theory also says that an agent will do a better job for the principal if the two parties' objectives are closely aligned. Thus if the realization of conditions hinges on cooperation and implementation by the agent, agent ownership of the project is essential. As Tirole (2002) puts it, in this case ownership is not a goal but a necessity.

IMF lending can also be cast in a principal-agent framework. In this case borrowing governments are the agents and the IMF—the delegated monitor of a revolving fund—is the principal. This principal-agent relationship is complex because of the

nature of the task, underlying loan contract, mandate and structure of the principal, and characteristics of the agents. This complexity, combined with the difficulty of specifying all contingencies in the contract, makes country ownership of programs all the more critical for their success.

The problem is that country ownership of IMF-supported programs is an elusive concept that is hard to pin down. Implicitly, it refers to a situation in which the policy content of the program is similar to what the country would have chosen in the absence of IMF involvement—because the country shares with the IMF both the objectives of the program and an understanding of the appropriate economic model linking those objectives to economic policies. In such a situation, the country owns the program in the sense that it is committed to its spirit rather than just complying with its letter. Thus the country will not deviate from its agreement with the IMF even if given the opportunity.

Because countries only borrow from the IMF when they face distress due to macro-economic or structural imbalances, sufficient safeguards are needed when providing access to IMF resources and to avoid moral hazard—requiring conditionality with some bite. Hence full country ownership is unlikely, and the real challenge is maximizing ownership in the context of conditionality. As noted, program conditionality is likely to place substantive constraints on the authorities' actions and use of funds. In addition, given that programs generally involve economic and social tradeoffs, perceptions may be created that conditionality does not take proper account of a country's circumstances, including its economic priorities, political conditions, culture, and traditions. This can—and sometimes does—lead to differing views on objectives, program strategies, and the pace of reform.

Ownership matters because it directly affects program implementation.³ Program agreements cannot envisage all the possible contingencies that could affect a program and specify in advance those actions that the authorities should take in response. When a country owns a program, decisions on such actions are likely to be made quickly and in support of the program, making it more likely that the program will succeed. In addition, ownership makes it easier to generate domestic political support for the program because it will likely be seen at least partly as an indigenous product, rather than a foreign imposition.

Ownership also matters for the catalytic role that IMF lending can play in increasing a country's access to foreign lenders. International flows of private capital have become increasingly important in recent decades, and a critical issue for borrowing in international markets is lenders' ability to exercise control rights. Foreign investors also confront moral hazard in teams: The payoffs to investments depend on the behavior and efforts of private borrowers as well as of the government of the country in which the private borrowers reside (Tirole 2002).

Firms investing across borders design appropriate covenants to mitigate moral hazard among private borrowers. Control rights not vested with the investor but

that affect borrower behavior are actually shared between the borrower and the government. Returns to a foreign investment depend on the environment created by government policies on domestic liquidity creation and management, tax and labor laws, and other institutional factors. When differences arise between a lender and a borrower or in times of distress, treatment of parties to a contract is crucially affected by public attitudes and policies toward law enforcement, bankruptcy, and corporate governance.

Government policies that reduce the amount of tradables or other internationally valuable collateral also hurt foreign investors. Such policies include taxing exports, failing to invest in infrastructure (which inhibits exports and tourism), depreciating the domestic currency when foreigners hold assets in that currency, depleting foreign exchange reserves, and creating incentives for currency and maturity mismatches that increase credit and default risks.

Whether such government moral hazard is important is an empirical issue and depends on the circumstances. It may be limited because governments lose power and credibility after a crisis, and in such cases IMF conditionality associated with adjustment programs places constraints on the authorities' policy choices (De Gregorio and others 1999). But it may be cause for concern: As in private firms, the threat of losing one's job after a crisis may prevent misbehavior—but it may also increase moral hazard by creating incentives to gamble for resurrection. Moreover, government actions that affect the mix of tradables and nontradables, and hence hurt foreign investors more than domestic ones, are less likely to generate adverse reactions from a country's population (Tirole 2002).

Hence country ownership of policies that reduce moral hazard related to foreign investment is likely to be important for a country's access to international capital markets. Such policies provide assurance to foreign investors that the government will not devalue their claims and, in the event of poor performance or adverse shocks to borrowers, will not inhibit the transfer of control rights to creditors. Of course, governments cannot relinquish control rights as easily as firms can, but that simply means that in their case the set of transferable rights will be more limited.

The feasibility of achieving a particular degree of ownership and determining when it has been achieved are problematic issues that vary by country. A complicating factor in assessing the degree of ownership is that most countries, especially democracies, have multiple stakeholders.

In pluralistic societies, does ownership refer to the views on program design and objectives held by the key ministers and central bank officials who negotiate the program with the IMF? Or does it refer to the views of the domestic bureaucracy that has to implement the program? Does it refer to the views of the parliament that has to approve the necessary legislation? Or to the views of civil society at large? If the views of civil society carry the most weight, how are they to be assessed and made to

influence program design and implementation in the face of competing interests? Thus ownership is intricately connected to trust in domestic institutions, effectiveness of political structures, and whether the government—negotiating on behalf of its citizens—has sufficient support to speak for a fair majority.

A widespread perception exists that program ownership by borrowing countries is insufficient. In recent years the number of program objectives has tended to increase because the IMF has taken on tasks that go beyond its traditional mandate of establishing macroeconomic and financial stability.⁴ In many programs the number of structural objectives has been expanded to facilitate transitions to market economies, integrate domestic economies with the world economy, diversify production and exports, develop financial sectors, and promote high-quality growth. In the 1990s these program goals were explicitly specified for transition economies and made prominent for Sub-Saharan African countries. In addition, in the Mexican (1994–95) and East Asian (1997–98) crises the key roles played by financial sector fragilities and corporate governance shortcomings expanded the list of goals. The problem is that country ownership is less likely when programs have too many objectives—because as the number of objectives increases, it becomes less likely that the authorities and the IMF will agree on the full range of objectives or on how they are to be attained.

Borrowing countries may be partly responsible for the lack of ownership. Some countries may be so eager for initial disbursements and for the catalytic role of IMF financing that they are willing to agree to programs without being convinced that the associated conditionality is appropriate. Such agreements have a greater chance of unraveling at critical decision points when it becomes clear that difficult policy measures are unlikely to be implemented. In private markets, if a lender has serious doubts about a borrower's intentions or is not provided sufficient collateral, the optimal course of action is not to lend. Given the IMF's cooperative structure, and the Samaritan's dilemma it faces, it is much more problematic for it to refuse to lend to a member in need (see Drazen and Fischer 1997).

IV. New Initiatives to Foster Greater Ownership

A number of proposals have recently been made to enhance country ownership of IMF-supported programs. Four such initiatives are considered here: encouraging countries to design their own programs (specifically in the context of PRSPs), streamlining structural conditionality, introducing flexibility in the timing of structural policy measures (floating tranche conditionality), and applying conditionality to outcomes rather than policies (outcomes-based conditionality).⁵ The IMF is already implementing some of these proposals (see IMF 2001a, 2001b, 2001c).

Encouraging Countries to Design Their Own Programs

The IMF could require or encourage borrowing countries to produce home-grown programs. In some cases this may be seen as forcing ownership, but program designs can be worked out cooperatively between the country and the IMF. If countries lack the expertise and capacity to develop their own programs, the IMF could provide technical assistance and training or encourage the authorities to hire independent technical advisers.

Although there are examples of home-grown programs, for several reasons this approach has generally not worked very well. First, countries often use overly optimistic assumptions when designing their programs. For example, countries may underestimate the extent of their difficulties and overestimate the potency of their policy instruments. Second, if the process of formulating a program—with its associated domestic political compromises—hardens the authorities' negotiating position with the IMF, it will likely cause significant delays in program negotiation. Third, countries may prefer to have the IMF prepare the program because they do not have (or choose not to have) domestic mechanisms for making decisions on difficult tradeoffs. This may require the IMF to force issues requiring decisions. Fourth, from a negotiating standpoint, countries may want to see the IMF's position before offering their own in a program document.

To increase program ownership by low-income countries, the IMF and World Bank recently began encouraging the production of PRSPs.⁶ The papers are intended to specify and detail a country's policies for reducing poverty (see Ames, Bhatt, and Plant 2002). PRSPs have three key elements:

- They are prepared by country authorities in consultation with various levels of government, local communities, civil society groups, donors, and multilateral agencies.
- They diagnose the country's poverty situation with the aim of identifying the main obstacles to raising incomes and welfare.
- They set goals for poverty reduction, define the immediate and long-term policy actions needed to achieve those goals, and design a system for monitoring progress.

Recent experience with PRSPs suggests that they could foster program ownership in several ways. (For a comprehensive review of the PRSP approach and its early experiences, see IMF 2002a, 2002b.) By creating a forum for dialogue—both within the government and among other stakeholders in society—that papers can identify the concerns and views of affected groups as well as policymakers. Thus they should enable policymakers and donors to better understand the different facets of poverty and the main concerns of poor people. In this regard, the challenge for governments is to strengthen democratic institutions and provide a voice to all domestic constitu-

encies. In many developing countries it cannot be assumed that nongovernmental organizations and civil society groups have the capacity to adequately participate in such a dialogue and engage in policy design, monitoring, and implementation.

PRSPs should also lead to better, more systematic collection of data and more effective monitoring. As a result, they can be used to manage expectations and set realistic goals, evaluate the tradeoffs and constraints involved, and develop priorities for policy action. Again, increased ownership will require ensuring that countries have the technical capacity to formulate and analyze different policy paths. In addition, the effectiveness of PRSPs will depend on how much they contribute to better governance and how policy strategies deal with corruption.

By offering a coherent strategy for attacking poverty, PRSPs could enhance coordination of program and nonprogram aid from multiple donors and multilateral agencies. Bulir and Hamann (2001) show that for many countries aid is highly volatile (more so than fiscal revenues), mildly procyclical, and not very predictable. As Svensson (2000) argues, when donors are unable to monitor reform efforts, aid disbursements are likely to be tied to economic performance—and hence procyclical.

Aid volatility and unpredictability can complicate expenditure management, especially for governments with deficient fiscal infrastructure, resulting in adjustments borne mostly by poorer and weaker segments of society. To the extent that PRSPs help coordinate donor efforts and reduce the cost of mobilizing and using aid, they could contribute to the success of IMF and World Bank programs—and so enhance country ownership.

For the IMF and World Bank, an important aspect of facilitating country production of poverty strategies will be coordinating advice and technical assistance to member countries. Such coordination will involve ensuring that the macroeconomic and financial conditionality attached to IMF lending is consistent with the sectoral and project-based conditionality (for structural and social policies) attached to World Bank lending. Such consistency will be achieved only if the IMF and the World Bank agree on their respective conditionalities, then use them in programs that respect—to the extent possible—the overall strategies articulated by member countries. A collaborative approach will enable countries to access and use effectively the IMF's Poverty Reduction and Growth Facility and the World Bank's Poverty Reduction Support Credit facility.

Streamlining Structural Conditionality

The past two decades have seen a major increase in structural conditions in IMF-supported programs (Goldstein 2000; IMF 2001c). The expansion of structural conditionality would appear to limit the scope for domestic policy choices, reducing country ownership. But there is considerable validity for the expansion (IMF 2001a, 2001b, 2001c). Even so, many structural reforms are of a microeconomic nature

and are likely to be more intrusive than macroeconomic policies. Country ownership of programs is essential for the design and implementation of these microeconomic measures because they have a different impact on various segments and vested interests in society.

It is widely felt that the IMF has gone too far with structural conditionality and overloaded programs with structural measures. Many structural reforms are not required to achieve macroeconomic stability. There is also no evidence that programs with more structural conditions have been more successful. In fact, programs with more structural conditions seem to have the same success rate as those with less (IMF 2001c).

Increased structural conditions pose two main dangers. First, they may reduce country ownership of programs and so impair their effectiveness. Second, a failure to implement structural reforms not critical for macroeconomic stability may undermine confidence in the overall program—possibly triggering reactions in domestic and international capital markets that could make overall program objectives harder to achieve.

It would be difficult and undesirable to turn back the clock and eliminate all structural conditions from programs. But careful thought should be given as to which structural reforms are essential to achieving a program's main objectives. These reforms will vary by country, but sharply pruning the list of structural conditions is possible without jeopardizing a program's success or the IMF's ability to be repaid. In other words, prioritizing or streamlining structural conditionality does not mean weakening overall conditionality.

The IMF has acknowledged that structural conditionality has expanded too much, and a major effort is under way to streamline it. IMF management recently defined broad principles for staff to use in determining the appropriate scope of structural conditionality in programs. In summary, the principles in the Interim Guidance Note on Streamlining Structural Conditionality (IMF 2001b:box 3) are:

- Structural reforms critical to achieving a program's macroeconomic objectives generally must be covered by IMF conditionality.
- Structural reforms relevant—but not critical—to a program's macroeconomic objectives and within the IMF's core areas of responsibility may be subject to conditionality.
- For structural reforms relevant to a program's macroeconomic objectives but neither critical nor in the IMF's core areas of responsibility, conditionality generally should not apply.

These principles represent an important shift by the IMF from comprehensive to parsimonious structural conditionality. Experiences with programs negotiated since the issuance of these principles will show whether this intention is being achieved.

Adopting Floating Tranche Conditionality

Performance criteria and structural benchmarks in IMF-supported programs have specific dates attached to them. Countries often find that rigid timetables for major structural reforms constrain their choices and strain their implementation capacity.⁷ Thus programs could be designed to allow for greater flexibility in the timing of structural reforms, increasing the scope for country ownership.

One way to achieve this goal is through floating tranche conditionality for structural reforms. Under this approach IMF loan disbursements would not be tied to specific dates. Instead, disbursements would be made available on completion of certain agreed-on reforms. This approach gives countries flexibility in the timing of program implementation.⁸ It also allows for disbursements associated with implementation of one part of a program to be unlinked from those associated with another part of a program.

The floating tranche approach could be used to divide conditionality into two parts. One part of IMF financing could be conditional on achieving the usual quantitative performance criteria under a predetermined schedule, whereas the other part could be dependent on implementing certain structural reforms at any time prior to the expiration of the program (and provided the macroeconomic program stays on track). In the floating tranche part the country would have control over when it undertook reforms and assurances that when it did it would receive related funding. The IMF would be protected because it would disburse funds only when reforms were undertaken.

The segmentation of conditionality would require decisions on the proportion of IMF financing subject to standard fixed tranche conditionality and that subject to floating tranche conditionality. Such decisions would be based on judgments about the relative importance of a program's different parts in achieving its overall objectives.

Not all structural reforms would be subject to floating tranche conditionality, because some are essential to macroeconomic improvements. For example, an independent central bank could be considered necessary to promote monetary stability, and a proper tax collection system might be needed to achieve fiscal discipline. The timing of such reforms could not be left open. In other words, in deciding which reforms are subject to fixed or floating tranche conditionality, the interdependence between structural measures and macroeconomic management would have to be taken into account. Final decisions would be made on a case-by-case basis and would be the result of negotiations and agreements between the IMF and the country authorities.

There is experience with a form of floating tranche conditionality in the context of the Higher Impact Adjustment Lending (HIAL) initiative, which was introduced by the World Bank in Sub-Saharan Africa in 1995. The tranching innovations under the initiative have two objectives: first, to give governments more freedom in the

timing of agreed reforms, thereby increasing ownership; second, to reduce pressure on the World Bank to disburse funds when conditions have not been met. These objectives are to be achieved through multiple but smaller tranches, increased disbursements after implementation of reforms, and the introduction of independent floating tranche arrangements.

Prior to the HIAL initiative, adjustment loans typically were disbursed in two tranches. HIAL introduced floating tranches, with single tranche operations as an alternative in special circumstances. Floating tranches are usually targeted at policy reforms in certain sectors and, in some cases, are in addition to regular tranches. Under HIAL, floating tranches have been applied to reforms involving the financial and banking sectors, the parastatal and public sectors, privatization, and civil service reform. A tranche is released only when the structural condition is met, regardless of when that happens. Of the 21 HIAL operations in 17 African countries in 1996–98, about two-thirds used the new tranching mechanisms.

In 1999 the World Bank's Operations Evaluation Department conducted an evaluation of these 21 operations (World Bank 1999). The study found that the HIAL initiative was associated with positive policy outcomes in terms of fiscal adjustment and exchange and interest rate policies. Moreover, the countries involved did better than nonparticipating comparator countries in increasing growth, lowering inflation, improving current account balances, stabilizing foreign exchange reserves, and achieving more sustainable debt paths. Even though HIAL programs differ in other ways from other World Bank programs, and the evaluation did not take into account the role of exogenous factors, the results provide some support for the use of floating tranches.

Basing Conditionality on Outcomes, Not Policies

Outcomes-based conditionality involves conditioning disbursements on the achievement of results rather than on the implementation of policies expected to eventually attain program objectives. Changing from policy-based conditionality to outcomes-based conditionality—leaving the choice of policies to the country authorities—has been advocated by, among others, Carlos Diaz Alejandro (1984:7):⁹ “I propose that the international community should return to the key rationale for conditionality, and negotiate with countries borrowing on concessional terms only regarding the balance-of-payments targets, leaving to the countries the decision as to what policy instruments should be employed to achieve them.”

Under the outcomes-based approach, IMF conditionality would focus on objectives rather than policy instruments and actions. Performance criteria for the disbursement of funds would be based on whether the targets for policy objectives were achieved by set dates. The policy objectives would be negotiated with the IMF, but the policy content of programs would largely be left up to country authorities.

This approach is not as radical as it might seem, because programs already define outcome variables as performance criteria. For example, IMF-supported programs include a floor on net international reserves as a performance criterion. Similarly, the adoption of an inflation targeting framework has made inflation one of the key variables monitored in the IMF's program for Brazil. Presumably, as more countries adopt the inflation targeting approach to monetary policy, more programs will follow the Brazilian model. Other variables that could be subject to outcomes-based conditionality include the trade balance, the current account, investment, growth, and so on.

In principle, there are two major benefits to this approach. First, the country authorities would be responsible for designing policies to achieve desired goals (as long as the policies are not prohibited by the IMF's Articles of Agreement). Hence the authorities would bear the risk of success or failure. This approach would enhance country ownership by requiring that the authorities and the IMF agree only on program objectives—and not necessarily on the mechanisms linking these objectives to specific policies.

Second, funds would be disbursed only on the attainment of certain goals, providing incentives for the country and the IMF to craft appropriate policies. IMF resources would be safeguarded because disbursements would depend on countries achieving the desired results. If policies did not have the envisioned outcomes, the country and the IMF would have to rethink the economic strategy.

Implementing this approach would undoubtedly present a few challenges. First, outcomes-based conditionality could lead to the backloading of funds that may be needed earlier to fill a temporary liquidity gap or to finance structural reforms. It would also lead to greater uncertainty for the country authorities about the availability of funds, because the agreed-on policies might not lead to the anticipated results.

Second, there may be significant lags in the reporting of data on outcomes, particularly for the real sector and for trade accounts. In addition, data on outcomes may be subject to frequent revisions, making timely monitoring and disbursements problematic.

Third, program outcomes are influenced not just by policies under the control of country authorities, but also by exogenous factors beyond their control. Though true in principle, it is not clear that exogenous shocks create serious problems for program projections. For example, a recent study of program projections by Musso and Phillips (2001) show that projections for growth, inflation, and international reserves were accurate relative to simple random-walk projections. But projections for current and capital accounts did not outperform projections from the random-walk model.

The more difficult question in this regard is, if exogenous factors force a program to go off course, to what extent should the authorities bear the risk of failure? Should

there be some assurance of IMF disbursement if it is judged that the authorities made a good effort to attain the goals? Thus even outcomes-based conditionality will require sifting the evidence to determine whether outcome targets were missed because of exogenous factors or because the authorities genuinely came up short—and if because of exogenous factors, whether there would be a case for waivers as there is under policy-based conditionality.

Outcomes-based conditionality also raises the question of whether and when it would be feasible for the IMF to disburse funds based on promises to achieve certain goals if it had not had any influence on the policy measures to attain them. In the private sector, such condition-free lending is made only to blue-chip clients or those with good collateral, high net worth, or both. Similarly, the IMF would likely provide such loans only to countries with strong records of economic performance and management, reputations for good governance, histories of paying debts as contracted, and those facing situations that are not too dire.

Otherwise, giving borrowing countries complete freedom in their choice of policy actions would not provide adequate safeguards for IMF resources. Even for the best clients, contract and loan covenants are likely to provide complete freedom for policy choices only as long as the client's capacity to repay is not impaired by endogenous or exogenous factors. To protect its resources, the IMF, like other lenders, will specify contingencies in the loan contract if the borrower's health deteriorates—for whatever reason.

Two points can be made in response to the problems raised regarding outcomes-based conditionality. First, even under outcomes-based conditionality funds will be disbursed in tranches. For example, a program to correct an imbalance in the balance of payments could take multiple steps to achieve its goal of attaining a comfortable level of international reserves. The first tranche could be disbursed based on a promise, but subsequent tranches would be released only after the country had achieved certain levels of international reserves. In fact, even the release of the first tranche could require some prior actions. Hence, by splitting monitoring and disbursement into smaller components, outcomes-based criteria can simultaneously provide sufficient safeguards and prevent excessive backloading of financing.

Second, like any creditor, the IMF would combine outcomes-based conditionality with a monitoring system so that if a borrower's position deteriorated sufficiently, it would intervene to contain the damage, take prompt corrective action, and try to change the borrower's strategy. So, as noted, there is likely to be *ex ante* ownership but *ex post* conditionality. That is, a program might have little bite initially, but stricter constraints would be imposed if certain events occurred. Because all possible contingencies cannot be specified *ex ante*, the IMF would design outcomes-based programs with the option of intervening should doing so be necessary to protect its resources.

V. Conclusion

This article has drawn on the well-established literature on agency theory and finance to argue that conditionality must apply to all IMF lending and that country ownership of IMF-supported programs is essential. Because ownership matters for the success of these programs, the IMF and country authorities share an incentive to create contracts that maximize ownership—subject to the safeguards that the IMF requires for its resources.

There is a widespread perception that countries often lack sufficient ownership of the programs they negotiate with the IMF. Here it is important to distinguish between countries in crisis and those with emerging imbalances. Crisis countries have less room for maneuver, and their problems require rapid responses. In such cases ownership may be less important—and, given the exceptional and extreme circumstances, agreements may be easier to reach. In other countries the extra time spent on negotiations may be a necessary price to pay for better program outcomes. In fact, different IMF facilities could cater to countries in different circumstances, and the scope and types of conditionality could differ depending on the facility used.

Ownership can be enhanced by limiting the objectives of IMF-supported programs—which would also allow for more focused conditionality. If a program's objectives are to establish and maintain macroeconomic and financial stability, the range of structural measures required by IMF conditionality would be narrower. In its capacity as policy adviser, the IMF can advise on the merits of various structural reforms. But in the context of programs, it should include only conditions that directly support macroeconomic objectives. Encouraging the domestic formulation of programs, selling programs to multiple country stakeholders, discussing different policy options, and increasing information flows would all help increase country ownership of programs.

So far the design of conditionality has focused on policy actions rather than outcomes. This article has argued that there is merit in shifting the emphasis toward outcomes-based conditionality and exploring the use of floating tranches, especially for structural reforms. Outcomes-based conditionality would increase country ownership by giving the authorities greater discretion and flexibility in choosing the policy mix and the timing of structural measures. This increased leeway in program implementation would be tied to explicit acknowledgment of ownership, improvements in data and reporting (to facilitate monitoring by external observers), and acceptance of responsibility for program outcomes.

For the IMF, outcomes-based conditionality would have to be combined with an agreed monitoring system for programs and establishment of rules for borrower behavior. Such rules would be applied uniformly and enforced through peer pressure and international norms—because the IMF has no recourse to legal action.

Although a good case can be made for incorporating outcomes-based conditionality in IMF lending, it is not an either/or matter. Programs would presumably combine

policy-based and outcomes-based conditionality. The balance would depend on the country's circumstances, preferences, and economic problems and on the accuracy with which different policy actions and outcomes can be monitored (Dixit 2000; Drazen and Fischer 1997). Programs with such a balance would align IMF conditionality more closely with country ownership—undoubtedly the shared goal of the IMF and its member countries.

Notes

Mohsin S. Khan is the director of the IMF Institute and Sunil Sharma is the chief of the Asian Division in the IMF Institute. The authors' e-mail addresses are mkhan@imf.org and ssharma@imf.org. The authors are grateful to Shanta Devarajan, Stanley Fischer, Morris Goldstein, Laurence Harris, Harold James, Vijay Kelkar, Peter Montiel, Jean Tirole, and several IMF colleagues for helpful discussions and comments. In addition, suggestions made by the article's referees helped sharpen the analysis. An earlier version of this article was presented at the third annual conference of the South Asia Network of Economic Research Institutes, held in New Delhi, India, on 28–31 August 2001. The views expressed in this article are the sole responsibility of the authors and do not necessarily reflect those of the IMF.

1. More formally, moral hazard can be defined as actions of economic agents maximizing their utility to the detriment of others in situations where they do not bear the full consequences of their actions due to uncertainty, asymmetric information, and incomplete or restricted contracts (see Kotowitz 1987; Stiglitz 2000).

2. Additional safeguards are provided by the fact that IMF claims are de facto senior to claims of other creditors and that funds are disbursed in tranches, conditional on the implementation of satisfactory policies (which are monitored) to correct the imbalances.

3. There is no direct empirical evidence on the link between ownership and IMF-supported programs. But some evidence on the importance of ownership for project lending is provided by the World Bank. The Bank's Operations Evaluation Department rates government commitment to each project (measuring in a sense the degree of ownership) using a variety of objective and subjective indicators. The relationship between project outcomes and government commitment turns out to be strongly positive (and statistically significant).

4. Feldstein (1998), for example, argues that the IMF has been too intrusive in its interventions and should ask two questions of each conditionality measure. First, is it needed to restore access to international capital markets? Second, would the IMF ask the same measure of a major industrial country if it had a program?

5. Other proposals not discussed here include preselecting countries eligible for IMF lending, developing policy options for country authorities to choose from, and investing time and effort in selling programs at home and abroad (see Khan and Sharma 2001).

6. Heavily indebted poor countries seeking debt relief and countries eligible to borrow from the IMF's Poverty Reduction and Growth Facility are required to produce a PRSP and have it approved by the Executive Boards of the IMF and the World Bank before seeking new program support.

7. In many cases rigid program schedules have posed serious problems for borrowing countries. One example involves the passage of laws: some legislatures have been surprised to learn that a program has committed them to not only a specific legislative agenda but also a set of deadlines under which the relevant legislative actions must be taken.

8. In principle, prior actions in IMF-supported programs can be thought of as a variant of floating tranche conditionality. A country agrees to undertake certain measures before the program (or program review) is discussed by the IMF's Executive Board. Thus the timing of the board meeting and the disbursement of funds depend on the prior actions having been taken. Reviews can also be considered a

form of floating tranche conditionality, because their completion (and accompanying disbursements) depends on agreed-on policy measures being taken.

9. Spraos (1986) makes a similar point but on the grounds that the links between outcomes and policy variables are too tenuous to make policy-based conditionality especially meaningful.

References

- Ames, Brian, Gita Bhatt, and Mark Plant. 2002. "Taking Stock of Poverty Reduction Efforts." *Finance and Development* (June):9–12.
- Anderson, Matthew S. 1966. *The Eastern Question: 1774–1923*. New York: St. Martin's Press.
- Boughton, James M. 2001. *Silent Revolution: The International Monetary Fund 1979–1989*. Washington, D.C.: International Monetary Fund.
- Buchanan, James M. 1975. "The Samaritan's Dilemma." In Edmund S. Phelps, ed., *Altruism, Morality and Economic Theory*. New York: Sage Foundation.
- Bulir, Ales and Javier Hamann. 2001. "How Volatile and Unpredictable Are Aid Flows, and What Are the Policy Implications?" IMF Working Paper 01/167. International Monetary Fund, Washington, D.C.
- De Gregorio, José, Barry Eichengreen, Takatoshi Ito, and Charles Wyplosz. 1999. "An Independent and Accountable IMF." Geneva Reports on the World Economy 1. International Center for Monetary and Banking Studies, Geneva.
- Diaz Alejandro, Carlos F. 1984. "IMF Conditionality: What Kind?" *PIDE Tidings* (January–February):7–9.
- Dixit, Avinash. 2000. "IMF Programs as Incentive Mechanisms." Working Paper, Princeton University, Princeton, N.J.
- Drazen, Allan and Stanley Fischer. 1997. "Conditionality and Selectivity in Lending by International Financial Institutions." IMF Working Paper. International Monetary Fund, Washington, D.C.
- Feldstein, Martin. 1998. "Refocusing the IMF." *Foreign Affairs* (March–April):20–33.
- Ferguson, Niall. 1998. *The House of Rothschild: Money's Prophets 1798–1849*. New York: Viking Press.
- Goldstein, Morris. 2000. "IMF Structural Conditionality: How Much Is Too Much?" Working Paper. Institute for International Economics, Washington, D.C.
- Greenbaum, Stuart I., and Anjan V. Thakor. 1995. *Contemporary Financial Intermediation*. Fort Worth, Tex.: Dryden Press.
- Gutián, Manuel. 1995. "Conditionality Past, Present, and Future." *IMF Staff Papers* 42(December):792–835.
- Haque, Nadeem Ul and Mohsin S. Khan. 2002. "Do IMF-Supported Programs Work? A Survey of the Cross-Country Empirical Evidence." In M. S. Khan, S. M. Nsouli, and C. H. Wong, eds., *Macroeconomic Management: Programs and Policies*. Washington, D.C.: International Monetary Fund.
- Holmström, Bengt. 1982. "Moral Hazard in Teams." *Bell Journal of Economics* 13:324–40.
- IMF. 2001a. "Conditionality in Fund-Supported Programs—Overview." Washington, D.C.: International Monetary Fund.
- . 2001b. "Conditionality in Fund-Supported Programs—Policy Issues." Washington, D.C.: International Monetary Fund.
- . 2001c. "Structural Conditionality in Fund-Supported Programs." Washington, D.C.: International Monetary Fund.
- . 2002a. "Review of the Poverty Reduction Strategy Paper Approach—Early Experience with Interim PRSPs and Full PRSPs." Washington, D.C.: International Monetary Fund.

- . 2002b. "Review of the Poverty Reduction Strategy Paper Approach—Main Findings and Issues for Discussion." Washington, D.C.: International Monetary Fund.
- James, Harold. 1996. *International Monetary Cooperation since Bretton Woods*. Washington, D.C.: International Monetary Fund.
- . 2001. *The End of Globalization: Lessons from the Great Depression*. Cambridge, Mass.: Harvard University Press.
- Kaplan, Steven N., and Per Stromberg. 2000. "Financial Contracting Theory Meets the Real World: An Empirical Analysis of Venture Capital Contracts." NBER Working Paper 7660. National Bureau for Economic Research, Cambridge, Mass.
- Khan, Mohsin S., and Sunil Sharma. 2001. "IMF Conditionality and Country Ownership of Programs." IMF Working Paper 01/142. International Monetary Fund, Washington, D.C.
- Killick, Tony. 1997. "Principals, Agents and the Failings of Conditionality." *Journal of International Development* 9:483–95.
- Kotowitz, Yehuda. 1987. "Moral Hazard." In John Eatwell, Murray Milgate, and Peter Newman, eds., *The New Palgrave: A Dictionary of Economics*. New York: Norton.
- Lindbeck, Assar and Jorgen Weibull. 1988. "Altruism and Time Consistency: The Economics of Fait Accompli." *Journal of Political Economy* 96(6):1165–82.
- Meltzer, A., C. F. Bergsten, C. W. Calomiris, T. Campbell, E. J. Feulner, W. L. Hoskins, R. L. Huber, M. H. Johnson, J. I. Levinson, J. D. Sachs, and E. E. Torres. 2000. *Results and Recommendations of the International Financial Institution Advisory Commission*. Final report to the U.S. Congress and Department of Treasury. Washington, D.C.
- Mishkin, Frederic. 1998. *The Economics of Money, Banking and Financial Markets*. New York: Addison-Wesley.
- Mussa, Michael and Miguel Savastano. 1999. "The IMF Approach to Economic Stabilization." IMF Working Paper 99/104. International Monetary Fund, Washington, D.C.
- Musso, Alberto and Steven Phillips. 2001. "Comparing Projections and Outcomes of IMF-Supported Programs." IMF Working Paper 01/45. International Monetary Fund, Washington, D.C.
- Polak, Jacques J. 1991. "The Changing Nature of IMF Conditionality." *Essays in International Finance* 184. Princeton University, Princeton, N.J.
- Santaella, Julio A. 1993. "Stabilization Programs and External Enforcement." *IMF Staff Papers* 40(September): 584–621.
- Spraos, John. 1986. "IMF Conditionality: Ineffectual, Inefficient, Mistargeted." *Essays in International Finance* 166. Princeton University, Princeton, N.J.
- Stiglitz, Joseph E. 2000. "The Contribution of the Economics of Information to Twentieth Century Economics." *Quarterly Journal of Economics* 115(4):1441–78.
- Svensson, Jakob. 2000. "When Is Foreign Aid Policy Credible? Aid Dependence and Conditionality." *Journal of Development Economics* 61(1):61–84.
- Tirole, Jean. 2002. *Financial Crises, Liquidity Provision and the International Monetary Financial System*. Princeton, N.J.: Princeton University Press.
- Williamson, John, ed. 1983. *Fund Conditionality*. Washington, D.C.: Institute for International Economics.
- World Bank. 1999. "High Impact Adjustment Lending (HIAL): Initial Evaluation." Operations Evaluation Department Report 19797. Washington, D.C.

Should Policy-Based Lending Still Involve Conditionality?

Stefan G. Koeberle

Traditional conditionality in policy-based lending is often criticized as being ineffective, intrusive, and corrosive. Disillusionment has led to proposals to replace ex ante conditionality with ex post conditionality and to focus on ownership, selectivity, and partnerships. This article reviews experiences with conditionality in the World Bank's policy-based lending and explores the benefits and drawbacks of various approaches. It argues that conditionality should play a central role in policy-based lending—but cannot substitute for country ownership and good policies. Moreover, an exclusive focus on conditionality based on ex ante commitments or ex post results may not be practical or useful for the Bank's policy-based lending. Thus a key recommendation is to use conditionality selectively, tailored to country circumstances. Indeed, an eclectic mix of traditional and new approaches is already being used—with programmatic policy-based lending offering a particularly promising way to reconcile the debate between the traditional ex ante approach and the aspirations of a results-based approach to conditionality.

Policy-based lending has inspired much debate, with much of the controversy revolving around the use of conditionality—the conditions attached to funds disbursed by international finance institutions. With research on aid effectiveness highlighting the importance of good policy and institutional environments and country ownership, the shortcomings of traditional conditionality have become increasingly apparent. Selectivity, partnerships, and outcome orientation have been suggested to address these shortcomings.

Is conditionality still useful? With donors focused on achieving the Millennium Development Goals, what type of conditionality remains useful and relevant? This article reviews experiences with conditionality in the World Bank's policy-based lending in light of the ongoing debate between the traditional approach to conditionality based on ex ante commitments to reforms and recent calls for a results-oriented

approach stemming from the literature on aid effectiveness. (Although the Bank sometimes imposes conditions in its other lending activities—such as investment lending or triggers defined in Country Assistance Strategies or programs for Heavily Indebted Poor Countries—here *conditionality* is defined more narrowly, as conditions attached to policy-based lending.)

The article finds that approaches to policy-based lending and conditionality have evolved considerably over the past decade and that conditionality has a central role to play when tailored to country circumstances. Country-specific conditionality requires a careful mix of traditional *ex ante* elements and new approaches. But this type of programmatic adjustment lending offers a promising way to reconcile tensions between country ownership and commitment to donors and between a medium-term strategy and flexibility.

Adjustment lending was originally conceived as a way of financing short-term balance of payments support—and to improve the policy environment for traditional project lending. Although the World Bank's Articles of Agreement envisage the provision of project finance as the Bank's primary activity, adjustment lending was introduced in 1980 under "special circumstances" allowed in article IV. (Though in fact, the Bank's first loans were program—that is, nonproject—loans.) But in response to changing borrower needs and broader reform agendas, new approaches to adjustment lending have evolved over the past two decades.

As a result, adjustment lending has become an important tool for supporting social, structural, and sector reforms, and is increasingly focused on long-run structural, social, and institutional issues. The narrow focus in the 1980s on achieving short-term stabilization and addressing distortions gave way in the 1990s to a more developmental perspective, with growing attention to reducing poverty, building institutions, and implementing complex social and structural reforms. This included an explicit focus on good governance, with strong support for public sector management reforms.

The Bank's adjustment lending now provides quick-disbursing policy-based financing based on actual or anticipated external financing gaps in the balance of payments or fiscal accounts (World Bank 1996). Adjustment lending is used to:

- Provide a cushion against economic shocks.
- Deliver external financing that generates local counterpart funds in support of government development programs.
- Promote policy reforms.

These three goals are often pursued concurrently. But there has been an evolution in the mix, with a shift away from short-term balance of payments support toward increased emphasis on medium-term external financing of government spending and support for social and structural reforms.

Over the years, World Bank policy-based lending has stirred substantial controversy. It has been the subject of extensive research and numerous internal Bank reviews

(World Bank 1986, 1990, 1992b, 1994; Thomas and others 1991). The Bank has been considering its experience with adjustment lending and conditionality as part of its periodic reviews of operational policies and lending instruments in support of country strategies and programs (World Bank 2001c, 2002c). The mixed performance of adjustment lending has been well documented (Mosley and others 1991). A more recent evaluation of developments in adjustment lending examined its appropriate use and the design of conditionality within the Bank's menu of lending instruments (World Bank 2001b). Based on several country studies, the Structural Adjustment Participatory Review Initiative drew broader conclusions on specific areas of policy reform, including privatization, agriculture, and the public sector (SAPRI 2002; World Bank 2001a).

This article first discusses critiques of traditional conditionality, along with evidence on the validity of such criticisms. It then assesses new approaches to conditionality in the context of World Bank experiences. After that the article discusses how suggestions for greater emphasis on country ownership and selectivity play out in real lending decisions. The article then reviews the implications of new approaches to conditionality, followed by an examination of a partnership approach to conditionality. The final section provides a summary of lessons on the use of conditionality.

Critiques of Traditional Conditionality

Conditionality is central to policy-based lending, linking financial support from donors to policy reforms considered critical for a country's economic and social development. The Bank memorandum introducing adjustment lending defines it as an agreement with a borrower to implement structural changes over a three- to five-year period, with financial support and technical assistance provided throughout (World Bank 1980).

Adjustment lending starts with the Bank and the borrowing government agreeing on needed policy reforms. The Bank then provides financing conditional on the implementation of specific reforms. This approach provides assurances to both parties—to the Bank that the reforms will be implemented and to the country that compliance will make financing available. Thus, conditionality also involves monitoring of whether country programs achieve their goals.

Conditionality has generated many contentious debates in the development community. The two biggest questions are: Does conditionality influence policy choices? Have the policies associated with adjustment lending led to better outcomes?

When assessing the effects of conditionality and adjustment lending, it is difficult (if not impossible) to devise appropriate counterfactuals and attribute outcomes to specific operations—especially because Bank adjustment lending often accounts for a small portion of government resources. To interpret results correctly, the effects of

adjustment programs must be separated from the often dire economic conditions that prompted them. It is also useful to distinguish World Bank adjustment lending from adjustment efforts supported by the International Monetary Fund (IMF), multilateral development banks, and other institutions. But methodological problems are compounded by the complexity of distinguishing the effects of different adjustment measures on different groups. For example, a currency devaluation may benefit poor rural farmers producing cash crops, but its inflationary effect may hurt poor urban workers.

Notwithstanding these empirical difficulties, assessments of conditionality have been mixed and inconsistent. Critiques generally fall into the following areas:

- *Efficacy and enforcement.* Traditional *ex ante* conditionality has been described as a failed instrument for promoting reform and growth—one that is useless and potentially damaging (Collier 1997; Easterly 2001). Its usefulness is undermined by difficulties in monitoring compliance and by incentives for donors to continue disbursing funds despite lax reform efforts. Because borrowers do not see the withholding of funds as a credible threat, they may not follow through on program commitments. The usefulness of conditionality is further diminished when donors engage in defensive lending—that is, grant new loans to help countries pay off old ones—regardless of whether the countries have implemented reforms.
- *Sustainability.* Another critique of conditionality questions the sustainability of externally induced reforms. Many analysts argue that once financing for a donor-supported adjustment program ends, reforms are often reversed or abandoned.
- *Process.* Conditionality is often seen as infringing on the sovereignty of borrowing countries. Some critics also question the legitimacy of the World Bank and the IMF, describing them as unelected and unrepresentative bodies that force countries to adopt policies that are not in their own best interests. In addition, there are concerns that agreements on conditionality are reached in nontransparent discussions between small groups of government officials and World Bank (and IMF) representatives without due consideration and participation by stakeholders, including civil society.
- *Content.* Some critics suggest that the main problem is the content of conditions, not conditionality itself. In this view, conditions are often based on best practices and theoretical fixes, with no objective investigation of their actual economic effectiveness and no consideration of political, social, cultural, and environmental features and limitations (Wood and Lockwood 1999). Policy-based lending is thus seen as the handmaiden of the supposedly malignant “Washington consensus.” For example, Milanovic (2003) asks why—despite numerous adjustment loans and IMF programs—Africa’s per capita income is the same as it was 20 years ago.
- *Proliferation of conditions.* Critics also charge that that recipient countries have been micromanaged by conditions that are too numerous, too detailed, and too

intrusive. Wood and Lockwood (1999) argue that more conditions have been placed on more areas of government policy, yet the central focus of reform programs has not changed. There is also concern that the number of conditions has increased as more bilateral donors have moved from project lending to policy-based lending in areas where they have an interest, including efforts to promote human rights, combat corruption, and build social capital.

Efficacy, Enforcement, Sustainability, and Process

Criticisms of conditionality in World Bank policy-based lending suggest an inconsistent perception of its influence on country policies. On one hand, there is widespread concern about its intrusiveness. On the other hand, attention is drawn to the limited effectiveness of traditional Bank adjustment lending.

Yet several country studies have found that conditionality has been useful, with reformers welcoming it and using the associated external commitment to push through reforms—especially when conditionality focuses on a few important measures to which governments were already committed. For instance, government officials in Ghana and Uganda welcomed conditionality that helped them identify, implement, and cement reforms (Devarajan and others 2001). Kenya, Tanzania, and Zambia offer less effective examples of the positive role of adjustment lending in supporting reforms (Burnside and Dollar 1997). But in other countries commitments to reform have been delayed, were not carried out, or were implemented but subsequently reversed.

Disappointing economic performance and uneven progress in policy reforms have been variously interpreted as failures of adjustment or as failures to adjust (World Bank 1994, 2001a). A clear verdict is thwarted by the serious identification problems inherent to evaluations of conditionality. Still, it is instructive to consider the regular evaluations of adjustment operations conducted by the World Bank's independent Operations Evaluation Department (OED). These evaluations assess project outcomes, likely sustainability, and institutional development impact. Projects are considered to have satisfactory outcomes if they achieve or exceed their main goals; in the case of adjustment operations, this includes the relevance of and compliance with conditionality (World Bank 1999a).

These ratings may be subject to methodological biases due to the problem of attributing outcomes to projects and given that evaluations are undertaken shortly after the funds are fully disbursed, before long-term results can be observed. Nevertheless, the evaluations indicate that most World Bank adjustment operations seem to meet their development objectives, are likely to be sustainable, and have significant impacts on institutional development—with significant improvements in the 1990s. OED outcome scores for adjustment lending increased from 60 percent satisfactory in the 1980s to 67 percent in fiscal 1990–94, then rose to 83 percent in fiscal 1999–2002.¹

The likely sustainability of adjustment programs also increased considerably, from 33 percent in the 1980s to 81 percent in fiscal 1999–2002, as did their institutional development impact, which rose from 21 percent to 48 percent. For all three indicators the findings for adjustment lending exceed those for investment lending. As will be discussed further, most *ex ante* conditions in World Bank adjustment loans have been met. Although most examples of involving civil society in the preparation or implementation of conditionality have emerged on an informal basis, recent adjustment programs based on the participatory processes involved in preparing Poverty Reduction Strategy Papers have increasingly recognized the importance of building consensus among stakeholders (World Bank 2001b). Thus, in contrast to the failures in efficacy, enforcement, sustainability, and process claimed by many critics, a systematic evaluation suggests that adjustment operations have on balance yielded substantial benefits for recipient countries.

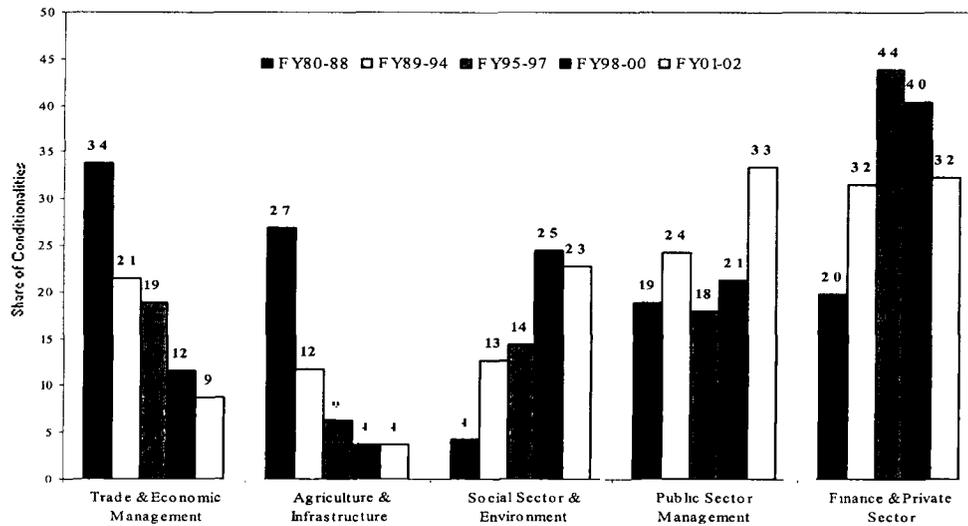
Content

Criticisms of conditionality often claim that a standard blueprint—presumably based on the Washington consensus—is used in all borrowing countries regardless of their circumstances (Wood and Lockwood 1999). Although the Bank's current operational policy contains prescriptive passages, this interpretation ignores the profound changes in World Bank policy-based lending over the past two decades (World Bank 2001b). The Bank's update of its operational policy on adjustment lending suggests refraining from using any blueprint for country-specific policy reforms (World Bank 2002b). Moreover, the validity of adhering to the Washington consensus has been called into question, including by the Bank's former chief economist (Stiglitz 1998).

The nature of conditionality in adjustment lending has evolved in line with the content of policy-based lending, reflecting the changing reform agenda in borrowing countries. Since being introduced in 1980 to help developing economies adjust their balance of payments after the 1979 oil shock, the Bank's policy-based lending has shifted from an early focus on supporting fiscal adjustment in response to external shocks to removing obstacles to growth and helping countries grow out of debt (figure 1). Until the early 1990s policy-based lending focused on adjusting relative prices that had been distorted by decades of import-substituting industrialization policies.

Today policy-based lending more often supports institutional reforms in public sector management and in the financial and social sectors. Many borrowing countries have moved beyond first-generation reforms involving the removal of economic distortions and are predominantly engaged in more complex reforms—building capacity and developing institutional infrastructure. As a result the share of policy conditions applied to agriculture and infrastructure reforms fell from 27 percent in

Figure 1. Conditions applied to World Bank adjustment loans by theme, fiscal 1980–2002 (percentage of all conditions applied)



Source: World Bank AICID.

fiscal 1980–88 to 4 percent in fiscal 2001–02. During the same period the share of conditions for public sector reforms increased from 19 percent to 33 percent, and for financial and private sector reforms from 20 percent to 32 percent. Across all sectors, the reforms supported by adjustment lending are increasingly long-term, institutional, and microeconomic.

Proliferation of Conditions

Disillusionment with traditional conditionality tends to focus on the number of conditions imposed, often calling them excessive (Wood and Lockwood 1999). There is no “right” number of conditions: according to the World Bank’s operational policy, a priori limits on the number of conditions are undesirable because they restrict the number of reforms that can be supported.

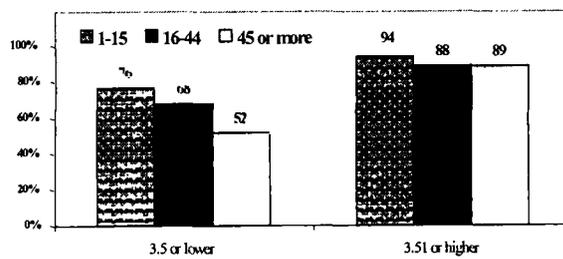
Moreover, determining actual conditions in adjustment loans is often not straightforward. Not only do policy agendas often contain a multitude of conditions, they also tend to mix reform measures, desired actions, and detailed processing steps. The average number of conditions in adjustment operations is still high, though it has fallen significantly—from 36 in the early 1990s to 21 in fiscal 2002.² More conditions tend to be negatively correlated with ratings for

adjustment loan outcomes—especially in countries with weak policies and institutions, as measured by the World Bank’s annual Country Performance and Institutional Assessment (CPIA). Thus country conditions matter in the design of conditionality.

At the same time, reflecting the growing focus on the long-run structural and institutional aspects of country reforms, conditionality has grown more complex (from 73 percent of adjustment operations in fiscal 1990–94 to 86 percent in fiscal 1999–2000, as rated by OED), demanding (from 82 percent of adjustment operations in fiscal 1990–94 to 93 percent in fiscal 1999–2000) and risky (from 72 percent in fiscal 1990–94 to 81 percent in fiscal 1999–2000) (World Bank 2001b).³ Adjustment operations considered complex also tended to have more waivers and lower ratings for outcomes and sustainability. These results can be largely explained by the fact that the number of conditions tends to be higher and complexity a greater challenge in countries with weak policy performance and institutional capacity, where adjustment lending is less successful. For example, in fiscal 1998–2000 adjustment operations in countries with low policy performance (CPIA ratings of 2.0–2.5) had an average of 39 binding conditions, compared with 21 in countries with high policy performance ratings (CPIA rating of 4.5 or higher).⁴

But outcomes highlight the ineffectiveness of efforts to address performance deficiencies and capacity limitations through a larger number of conditions. Adjustment operations were less successful in countries with weak policy performance subject to more conditions, whereas countries with stronger performance did well regardless of the number of conditions (figure 2). Thus a country’s policies and institutions are typically more important than specific design features—such as the number of conditions—in determining the success of a program.

Figure 2. Adjustment Operations with Satisfactory Outcomes by Number of Conditions and Country Performance, Fiscal 1990–2002 (percent)



Source: World Bank Operations Policy and Country Services calculations based on World Bank AICID and OED data.

Suggested New Approaches to Conditionality

Because traditional conditionality has not been entirely successful, many critics have wondered whether it should be dropped altogether. Various new approaches have been suggested to address the inherent tensions in conditionality.

- *Increasing country ownership and selectivity.* One issue involves the potential conflict between country ownership and a lender's due diligence and enforcement of loan conditions. The importance of strong policies and institutions for successful reform, as highlighted by the literature on aid effectiveness, has also given rise to calls for selectivity in favor of better-performing countries.
- *Improving the design of conditionality.* Because the traditional approach to policy-based lending—using *ex ante* conditionality based on promised actions—has been perceived as useless (at best) and even corrosive, an approach based on reputation and results has often been advocated.
- *Strengthening partnerships on conditionality.* A renewed focus on partnerships and coordination—between recipients and donors and among donors—has gained currency in response to traditional conditionality perceived as overly intrusive.

Increasing Country Ownership and Selectivity

Many debates on conditionality have focused on the importance of country ownership—that is, commitment to aid-supported reforms by country authorities and a majority of domestic stakeholders. Country ownership has always been a central concept in development aid, on the assumption that it makes the policy and institutional changes associated with lending operations more likely to be implemented even in the face of political opposition. But in recent years the literature on aid effectiveness has increased the emphasis on ownership, driven by the proposition that donors can only advise and support but not buy or induce economic reforms (World Bank 1998a). Thus government willingness to reform—ideally supported by a broad consensus among members of civil society—is considered essential to successful adjustment programs.

The World Bank has long recognized the importance of country ownership and readiness to reform as critical factors for effective policy reforms and sustained development (Johnson and Wasty 1993). In light of the past two decades of policy-based lending, the Bank now shares the view that conditionality can reinforce country ownership—but not substitute for it (World Bank 2001b).

More recently, the Bank has made concerted efforts to form results-oriented partnerships for development to reflect three main lessons highlighted at the 2002 U.N. Conference on Financing for Development (held in Monterrey, Mexico). First, good development outcomes require appropriate policies and institutions. Second, sustained

development progress requires that policies and institutions be country-owned and country-specific. Third, when these conditions are in place, development assistance can be highly effective (World Bank 2002a).

Country ownership and reform readiness are difficult to assess. Several conceptual frameworks have been developed to assess the level and quality of country ownership:

- *Leadership analysis* assesses senior policymakers in terms of their initiative in formulating and implementing reforms, their level of intellectual conviction, their expression of political will, and their efforts to build consensus among various constituencies (Johnson and Wasty 1993). Leadership is important for successful reform, but senior policymakers may underestimate difficulties in securing support from other political actors and sustaining institutional efforts.
- *Stakeholder analysis* focuses on understanding the power relationships, influences, and interests of stakeholders affected by policy reforms, including those in government (Heaver and Israel 1986). This analysis also captures the extent to which stakeholders can make their voices heard, participate in decisionmaking, reach consensus, and accept short-term costs and uncertainties in the distribution of benefits in exchange for long-term gains.
- *Reform readiness analysis* captures the commitment and performance of key policymakers and interest groups and so requires detailed knowledge of a country's political situation (Haggarty and Matuda 1999). Based on findings that successful reforms must be politically desirable, feasible, and sustainable, this approach examines institutional arrangements for policymaking and analyzes the political rationale for particular policy decisions.

Each of these approaches has merits and limitations (Morrow 1999). The challenge is identifying robust indicators for judging whether policy-based lending is appropriate in a specific country situation. These might include the authorities' willingness to prepare an action plan outlining the government's reform intentions and the extent of consultations with and participation of civil society in designing and implementing reforms.

A country's track record of policy performance is among the most robust predictors of whether reforms will be implemented. But some countries with weak track records have managed to turn themselves around fairly quickly. Such cases require a judgment that the risks of failure are outweighed by the potential rewards of continued engagement.

Decisions on policy-based lending cannot overlook the tensions between limited country ownership and the use of conditionality to ensure that reform objectives are achieved. If a country's commitment or implementation capacity is weak, conditionality is unlikely to be effective. By itself conditionality cannot lead to the adoption of better policies if there is no consensus for reform. When there is commitment and capacity to reform, adjustment lending can accelerate, broaden, and deepen it, enhancing its

impact and hopefully contributing to growth and development. But in the absence of such commitment, adjustment lending may fail to support improvements in policies and institutions—and indeed may contribute to delays in reform and leave the country burdened with debt.

One way to bolster country ownership is to forgo *ex ante* conditionality in favor of *ex post* allocations of policy-based loans based on the policies a country adopts. Although successful adjustment can deliver significant long-term benefits, research on aid effectiveness suggests that these benefits can be realized only if a country's policy environment is favorable (World Bank 1998a). This suggests that policy-based lending should focus on countries where it is most likely to be effective—those with good policies and high poverty.

Empirical studies emphasize that policy changes are driven primarily by domestic political economy, not by foreign assistance or policy-based lending (Devarajan and others 2001). Alesina and Dollar (1998) find that there is no tendency for surges in finance to lead to policy reforms. In fact, policy is generally quite persistent, and sharp policy changes are the exception, not the rule.

Greater selectivity in World Bank lending, based on country performance, is among the main suggestions for increasing aid effectiveness (World Bank 1998a). Taking that argument further, the Meltzer commission on international financial institutions advocated limiting policy-based lending to countries with good policies (Meltzer and others 2000). Despite several obvious shortcomings, these proposals have been influential.

First, although many advocates see it as a way of avoiding traditional *ex ante* conditionality, selectivity based on judgments of country performance may implicitly result in similar bargaining between the Bank and borrowing countries. Unless selectivity is based on a few measurable, objective indicators (such as spending allocations), assessing a country's performance is inherently subjective—particularly if it involves judging the relevance and effectiveness of a country's policy choices.

Second, selectivity would not eliminate the tensions between country ownership and donor intrusiveness. Instead the bargaining process would shift from conditionality to policy assessments.

Finally, there is no reasonable basis to objectively distinguish countries with good policies (where policy-based lending would make sense) from countries with bad policies (where it would not). Most of the Bank's borrowing countries fall between these two extremes, with mixed performance resulting in rather subjective judgments and lending decisions. Moreover, the countries in this heterogeneous middle ground are home to most of the world's poor people. If selectivity meant that adjustment loans went only to the small handful of countries with unquestionably good policy performance, these poor people would not benefit from any policy-based lending, however ineffective.

Aid allocations based on country performance are not a novel idea for the Bank. It already uses *CPIA* ratings in determining the relative country financing shares from

the International Development Association. In addition, its Country Assistance Strategies include different lending scenarios based on triggers typically driven by country performance. Moreover, in recent years considerable selectivity has emerged in decisions on adjustment lending.

Since the mid-1990s most—though not all—adjustment lending has gone to countries with above-average policy performance, for sectors with good track records, and for issues with adequate supporting analytical work. In fiscal 1996–2002 borrowers with above-average CPIA ratings received 72 percent of adjustment loans by volume (amount) and 57 percent by number.⁵ When adjustment lending has gone to countries with poor track records, assessments of potential risks and rewards have typically argued that such loans were needed to exploit brief windows of opportunity for long-overdue reforms. But some less successful examples—such as adjustment lending to the Russian Federation shortly before it defaulted on its foreign debt in 1998—show that despite the best intentions, it is extremely difficult to assess such situations.

Improving the Design of Conditionality

The contrast between traditional *ex ante* conditionality and a results-driven approach influences the design of conditionality. Some analysts have called for a results-oriented approach to conditionality based solely on outcomes and for the use of different loan disbursement procedures to align conditionality with country performance rather than promises.

Outcomes

Disillusionment with *ex ante* conditionality has led to suggestions that loan disbursements be linked to outcomes rather than policies. It is argued that specifying the desired results of reforms rather than policies on how to achieve them would signal that donors are concerned with the destination rather than the journey. Giving country authorities more discretion in achieving agreed outcomes is an attractive proposition that could eliminate intrusive donor micromanagement of policy and institutional reforms. Freed from the principal-agent problem of bargaining with donors, countries would have greater scope to define their policies and so have greater ownership.

But linking disbursements to outcomes creates several problems (Gunning 2000): First, outcomes result from many factors other than policy choices. Moreover, the link between government actions and outcomes is often not clear—as exemplified by the complex relationship between social spending and social outcomes (Filmer and others 2000; Thomas and others 2000). Thus, countries might suffer if disbursements

are withheld based on inadequate outcomes beyond the government's control. Collier and others (1997) suggest correcting for this bias by identifying important determinants of outcomes beyond the government's control (such as landlockedness and ethnolinguistic fractionalization). But this approach is limited to a few indicators and presupposes a good understanding of determinants and available data.

Second, it is harder to monitor outcomes than policies, and data on outcomes are often not available or are fraught with methodological problems. Finally, even if there is a clear link between government actions and outcomes and data on outcomes are available, outcomes typically change slowly and can be monitored only over the medium term—implying that current governments would be held responsible for the actions of their predecessors. Linking disbursements to outcomes would also be impractical for policy-based lending, which usually involves disbursements over periods of less than three years.

Outcome-based conditionality has been tested in a few countries. Under the leadership of the European Commission, the Special Programme of Assistance for Africa—sponsored by a group of multilateral and bilateral donors—experimented with a pilot approach in Burkina Faso that linked the release of funds to program outcomes rather than policy reforms. This effort involved joint evaluations by all donors of overall results rather than individual conditions (Emblad and Gilles 2000). Preliminary evaluations suggest that this approach increased country ownership and donor coordination (SPA Task Team on Contractual Relationships and Selectivity 2000), with implementation issues revolving around the choice of outcome indicators, realistic targets given time lags, and the quality of data (European Commission 2001).

In World Bank policy-based lending, conditionality based on outputs and outcomes has been used increasingly in recent years. Such conditionality does provide more flexibility to countries and makes conditionality more effective—provided outcome-related benchmarks are quantifiable, narrowly defined, and complemented by action-based conditions where appropriate. Benchmarks should also be clearly linked to measures over which policymakers have discretion to reduce borrower uncertainty about the release of financing. These indicators should be agreed at the outset of lending programs, with a clear understanding of which ones (primarily those related to policy actions) are within government control.

Phasing and Tranching

Debates on the best approach to conditionality also involve the timing and phasing of performance benchmarks supported by policy-based lending. When adjustment loans provide fast-disbursing financial support in response to economic distress, they tend to focus on problems amenable to short-term solutions—such as stroke-of-the-pen reforms of tariffs and regulations.

But when policy-based lending supports complex structural reforms in areas such as public sector management, privatization, and the financial sector, there is a tension between the rapid provision of financing and the often slower pace of intensive reforms. To reflect these timing considerations, adjustment loans are disbursed in tranches—that is, portions of funding are withheld until certain conditions are met. Tranches have typically been designed based on ex ante conditionality, though in principle they could be used for ex post approaches. Among 452 World Bank adjustment loans approved in fiscal 1990–2002, 85 percent fully disbursed all tranches as planned (World Bank Adjustment Lending and Conditionality Implementation Database, ALCID). Thus, the majority of operations did not experience waivers, restructuring, or cancellations, suggesting that a common criticism of traditional conditionality—that it leads to frequent noncompliance with policy conditions—has limited validity.

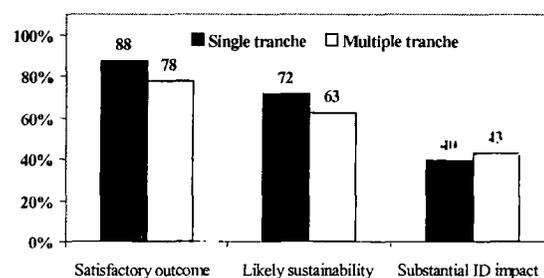
Still, appropriate phasing of policy-based lending to support countries' growing focus on complex medium-term structural and institutional reforms involves difficult choices among various options.

- *Multiple tranches.* The World Bank's traditional approach using ex ante conditionality has typically been based on multiple tranches to support reform commitments as they are achieved in a single operation. Multiyear and multitranche operations have been used to demonstrate country commitment and to lock in reforms through conditionality covering future actions. It is largely this type of ex ante conditionality that has given rise to charges of noncompliance, corrosion of sovereignty, and low efficacy. But in several countries this approach has been welcomed by reformers, strengthening their hand against vested interests, or served as a signaling device (Devarajan and others 2001). In principle, multiple tranches can also clarify future commitments when a country's track record is weak. But in practice, multiyear and multitranche operations account for the bulk of implementation difficulties that occasionally hamper adjustment programs. For example, during fiscal 1990–99, 39 percent of all approved two-tranche operations and 63 percent of three-tranche operations involved waivers or cancellations (World Bank ALCID). Such problems are especially pronounced in multiyear operations due to the challenge of setting out during negotiations a detailed policy plan covering actions to be taken a few years later.
- *Floating tranches.* The Bank has increasingly used floating tranches—that is, tranches disbursed when specific conditions are fulfilled rather than according to a set schedule. This approach can be appropriate for clearly defined discrete reform actions with uncertain timing, such as privatization of public enterprises. Preliminary evidence suggests that floating tranches for different parts of a reform program can increase the credibility of loan conditionality. Floating tranches generate complementarities and synergies among reform measures

and avoid delays in specific reforms holding up progress in the overall program. In addition, OED evaluations of the Higher Impact Adjustment Lending initiative in Sub-Saharan Africa have found that floating tranches provided considerable flexibility in timing and increased country ownership—especially when combined with fewer but well-focused conditions (World Bank 1998b, 1999b).

- *Single tranches.* Single-tranche operations have become more common. During fiscal 1998–2000 the Bank approved 36 single-tranche adjustment operations worth US\$12.8 billion, and by fiscal 2000 single-tranche loans accounted for nearly 40 percent of the number and volume of adjustment operations. Experiences with single-tranche operations have been encouraging, with OED evaluations for fiscal 1990–2002 showing higher ratings than for multitranche operations in terms of satisfactory outcomes (88 percent compared with 78 percent) and likely sustainability (72 percent and 63 percent; figure 3). (The higher ratings may partially reflect the wider use of single-tranche operations in countries with better policies and less need for conditionality.) Although single-tranche operations received slightly lower ratings in terms of their impact on institutional development, this may reflect their shorter disbursement periods. Indeed, concern about the sustainability of single-tranche operations is the main reason for the Bank’s operational policy provision that such operations may be appropriate in environments of high uncertainty—provided there is adequate ex ante conditionality and a satisfactory medium-term program in which the operations can be evaluated (World Bank 1996). Again, country conditions matter for effective conditionality design. In countries with below-average policy and institutional performance, single-tranche operations are associated with better outcomes and fewer waivers, whereas higher-performing countries achieve good outcomes regardless of the number of tranches.

Figure 3. Number of Tranches and Quality of Adjustment Loans, Fiscal 1990–2002 (percent)



Source: World Bank Operations Policy and Country Services calculations based on World Bank AICD and OED data.

- *Programmatic adjustment lending.* Over the past few years the Bank has increasingly used a programmatic approach for its policy-based lending. This approach involves a series of single-tranche operations that are subsequently presented to the Bank's Board of Executive Directors, with a medium-term framework specified at the outset—including completed prior actions, monitorable progress indicators, and expected prior actions for subsequent operations. To the extent possible, programmatic approaches align disbursements with the borrowing country's annual budget cycle. By building on completed actions specified and agreed to in advance (instead of future promises), a programmatic approach combines country ownership with systematic reform implementation. Programmatic lending is usually used to support complex medium-term institutional reforms. Poverty Reduction Support Credits are emerging as a vehicle that incorporates these principles, basing social and structural reforms in IDA countries on the poverty reduction strategies articulated in these countries' Poverty Reduction Strategy Papers (World Bank 2001b)

Country experiences suggest that an exclusive focus on conditionality based on ex ante commitments or ex post results may not be practical or useful for World Bank policy lending. The tranching and phasing options described are not appropriate in every country or sector. Traditional multiyear, multitranche designs may still be useful, especially for supporting in-depth sector reforms. Floating tranches can help with discrete reforms with uncertain timing without holding up the rest of the program. For strong performers with good track records—such as Brazil, Latvia, Mexico, and Uganda—the programmatic approach has been used to provide sustained, flexible, and predictable support to credible medium-term reform programs. The programmatic approach has also been useful in countries with sound reform programs but weaker capacity or track records or in countries emerging from crisis or instability—such as Jamaica, Peru, Turkey, and Ukraine. Early experience suggests that the programmatic approach is well suited for its goals of fostering country ownership, providing reliable financial support for successful medium-term programs, and accommodating the uncertainties inherent in medium-term reforms (World Bank forthcoming).

Monitoring Results

Debates on conditionality implicitly emphasize the importance of monitoring its development impact, using specific indicators to judge its effect on the borrowing country's compliance (in the case of ex ante conditionality) or performance (in the case of a results-based approach). Borrower implementation and Bank supervision of lending operations have been key issues for the World Bank—especially since the Wapenhans Report stressed their importance (World Bank 1992a). Past efforts to

strengthen monitoring and evaluation of the Bank's loan portfolio focused on investment lending (World Bank 2000).

For policy-based lending, monitoring has typically focused more on compliance with ex ante conditionality than on progress, outcomes, and poverty impacts. All programs supported by policy-based lending, whether based on an ex ante or an ex post approach, involve implicit assumptions about the expected effects of certain actions on economic performance and incentives, social conditions, poverty reduction, and the environment. But it is difficult to define and apply performance indicators that compare actual and expected outcomes and allow for program corrections. Linking policy measures supported by conditionality and their associated financing to specific economic and social outcomes entails several problems:

- Reform programs have complex, economywide repercussions.
- Poverty reduction and economic performance are influenced by many other factors.
- Reliable and timely data are often not available.
- Changes in countrywide economic, social, and environmental trends may be realized only with a considerable lag.

Evaluations of policy-based lending are further complicated by the difficulty of specifying counterfactuals—that is, what would have happened without the operations.

The Millennium Development Goals—the product of the U.N. Millennium Declaration adopted by the international community in 2000—pose a new challenge for increasing the focus on results (IMF and others 2000). Tying adjustment lending to development outcomes is a critical part of the Bank's commitment to align country and sector strategies with the eight development goals (World Bank 2001d). These goals provide an ambitious yardstick against which to measure the contribution of conditionality to alleviating poverty, increasing school enrollments, reducing child and maternal mortality, expanding access to reproductive health services, eliminating gender disparities, and improving environmental management.

Aligning conditionality with the Millennium Development Goals requires realistic timeframes, intermediate goals, and measures differentiated by countries and regions. Country programs for policy and institutional reforms should monitor clear performance targets over the medium term while maintaining scope for adapting reform efforts to changing country circumstances. Programmatic adjustment lending offers a promising way to do both.

Even so, it has been challenging to develop a road map for policy-based operations that allows meaningful assessment of reform progress and goes beyond mere compliance with policy conditions, whether ex ante or ex post. Policy reform plans must spell out a practical framework that links strategies to results using monitorable progress indicators. (For programmatic operations, such indicators should complement the actions selected as expected prior actions—or “triggers”—for subsequent operations.) It is no trivial task to define performance indicators precise enough to guide

implementation and enable unambiguous monitoring of results—yet pragmatic and flexible enough to allow for the inherent unpredictability of institutional reforms over the medium term (see World Bank 2001b for a discussion of performance benchmarks).

Regardless of whether an *ex ante* or *ex post* approach is chosen, conditionality will yield results only if its design reflects a country's institutional capacity. Clear monitoring arrangements are integral to effective implementation, including that by local governments and civil society where appropriate. In addition, governments must build domestic capacity for monitoring economic, social, poverty, and environmental outcomes.⁶

Strengthening Partnerships on Conditionality

Increased partnerships and collaboration between donors and borrowers and among donors have been part of the response to the claim that traditional *ex ante* conditionality is overly intrusive. It is hoped that transparent partnerships between donors and countries will encourage true country ownership of reforms. The most sweeping proposal in this area is to channel all development assistance through a common pool instead of financing individual programs and projects—augmenting the general budgets of poor countries less obtrusively (Kanbur and Sandler 1999).

The World Bank supports coordination and harmonization efforts, which attempt to combine partnerships with a focus on results—most recently in the Rome Declaration on Harmonization (World Bank 2003). The Bank's Comprehensive Development Framework advocates a long-term holistic approach, with borrowing countries owning and directing their development strategies and stronger partnerships among governments, donors, civil society, the private sector, and other development stakeholders in implementing the strategies. The goal is to make aid more effective through coordinated support for country programs and to benefit from synergies and avoid duplication among different development partners, including the IMF, other multilateral development banks, and bilateral donors.

Applying these principles to conditionality requires aligning the design and implementation of policy-based lending supported by the Bank, the IMF, and other multilateral development banks, which have traditionally provided the bulk of policy-based lending. Efforts to coordinate policy-based lending aim to strengthen government ownership by reducing the scope and number of conditions, making policy-based lending more effective in supporting poverty reduction efforts, reducing burdens on country capacity, and enhancing the predictability of aid flows. In addition, as bilateral donors shift away from projects toward direct budget assistance supporting poverty reduction strategies, their support needs to be coordinated with country processes and other donors.

One risk of these partnership approaches to conditionality is that they tend to ignore tensions between stakeholders with different incentives. It is in the choice of conditionality where potential conflicts between ownership and intrusiveness are played out, when the agenda of donors diverges from that of reluctant reformers. Because few countries are monolithic, policy choices reflect the various influences that domestic stakeholders and competing vested interests have on policymakers. The political economy of borrowers might also create incentives for policy reversals after an adjustment program has been concluded.

Moreover, because the effectiveness of donors such as the World Bank is measured as much by lending volumes as by program outcomes, such donors tend to overestimate the likelihood of successful reforms in borrowing countries. In addition, different donors may have different agendas, potentially leading to conflicting incentives. If these complexities are not addressed, the partnership approach to conditionality risks generating similar disenchantment as traditional conditionality.

Coordination

Efforts to coordinate conditionality among multilateral development institutions have focused on the World Bank and the IMF—the largest providers of policy-based lending. The Bank and the IMF have distinct but complementary responsibilities and expertise for supporting member countries' adjustment programs. Both institutions have long had a framework for collaboration to help ensure that their advice is consistent and fully exploits each institution's expertise and financial resources—as reflected in the so-called joint guidelines (World Bank and IMF 1998).

Poverty Reduction Strategy Papers and the Heavily Indebted Poor Countries initiative have contributed to more systematic arrangements for Bank and IMF collaboration in many low-income countries, with each institution's conditionality focused on its areas of primary responsibility. Extending similar principles to other countries, the Bank and the IMF have been strengthening their collaboration in supporting country development efforts through more coherent, streamlined conditionality (World Bank and IMF 2001). Based on the premise that increased collaboration on conditionality would strengthen program designs, a “lead agency” concept was introduced to deal with specific policy issues, along with systematic information sharing and monitoring. Enhanced collaboration is being operationalized through increased interaction between the staff of the two institutions and through transparent reporting in board documents of each institution's views on borrowing countries' reform priorities, program conditionality, and progress in program implementation (World Bank and IMF 2002).

Among the forces driving increased Bank and IMF collaboration are ongoing efforts by the IMF to streamline its conditionality and increase country ownership of the programs it supports. In the past the extended use of structural benchmarks in

IMF-supported programs contributed to significant overlap in the structural conditions of IMF and World Bank-supported programs (World Bank 2001b; IMF 2001a). By contrast, conditionality in the Bank's policy-based lending has tended to focus on structural reforms in areas within its mandate, with a general provision that countries receiving adjustment loans maintain an adequate macroeconomic framework. The main aim of the IMF's review of its conditionality was to focus it on measures critical to monitoring and achieving the macroeconomic objectives of IMF-supported programs and to apply it sparingly outside the IMF's core areas of responsibility (IMF 2001a, 2001b, 2002).

The boards of the IMF and the World Bank stressed that the streamlining of IMF conditionality should result in an overall reduction in the conditions imposed on borrowing countries—and not just shift them from one institution to the other. In addition, enhanced reporting on conditionality to the boards of both institutions is designed to ensure that areas no longer covered by IMF conditionality will be adequately addressed by the Bank or other institutions.

Success in both institutions' efforts to streamline and focus conditionality hinges on effective collaboration. The staffs of both institutions must develop a shared vision of their support for each borrowing country, with a clear division of labor based on each institution's areas of expertise and frequent dialogue and information sharing.

Budget Support

The problems of donor-driven project proliferation are well known, including the stress it imposes on country capacity, the potential substitution of donor-driven agendas for country ownership, and the fragmentation of country budget processes. In response, donors (especially bilateral donors) have sought to enhance the effectiveness of their aid programs by replacing or supplementing traditional project financing with direct budget support—particularly in low-income countries developing Poverty Reduction Strategy Papers. Recent recipients of such support include Ghana, Tanzania, and Uganda.

Whether intended or not, this type of aid resembles the policy-based lending provided by the World Bank, the IMF, and other multilateral development institutions. Although the objective is typically to provide predictable and less intrusive financing and greater country ownership, by its nature budget support implies that conditionality explicitly or implicitly guides donor decisions.

The shift from project to nonproject financing may raise a host of new issues, including concerns about:

- Poverty Reduction Strategy Papers that are not specific on programs, action plans, and output indicators, with weak or no links to the budget cycle or medium-term spending framework.

- Discrepancies between poverty reduction strategies and donor programs.
- Contradictions or overlap between the budget support provided by different donors.

Moreover, instead of facilitating the alignment of donor support with the borrowing country's budget cycle, there is a risk that discordant donor priorities or processes may excessively strain country capacity. Even the goal of predictable resource flows could be undermined if donor decisions on disbursements are based on different performance criteria.

These issues call for much closer coordination of policy-based lending, with explicit recognition of the need for a systematic approach to conditionality. Such efforts may involve coordinating the timing, content, and process of donor negotiations with governments; assigning clearer divisions of labor or even lead agencies for specific areas; and focusing conditionality on a small set of country priorities and plans developed by the borrowing country. Such approaches would make it possible to better integrate sector programs and capacity building with countries' regular budget processes. Although donors would maintain distinct accountability for their decisions, they would aim to increase the predictability of resource flows through coordinated assessments of borrowing countries' performance against agreed benchmarks.

Conclusion

Development financing in support of country policies has become more important as donors increasingly recognize the need for strong country policies and institutions and shift away from narrow project financing. Yet conditionality remains controversial and is often considered intrusive, ineffective, or even harmful. Disillusionment with traditional *ex ante* conditionality has led to proposals for more results-orientated approaches focused on ownership, selectivity, and partnerships.

These approaches have merits and drawbacks in the reality of World Bank lending decisions. Moreover, many of their elements have already been incorporated into the Bank's policy-based lending—and the improving quality of adjustment lending suggests that Bank conditionality reflects some of the lessons from the past two decades of experience. In the future these approaches should be made more systematic and transparent by spelling out how the policies and institutional reforms they support will contribute to the achievement of country objectives. In addition, monitoring of the poverty outcomes of policy-based lending should include indicators linking conditionality to progress toward the Millennium Development Goals wherever possible. Developing such indicators will require concerted diagnostic and analytical work.

Overall, policy-based lending remains a useful tool for supporting government reforms. Conditionality is integral to policy-based lending—whether explicitly

included *ex ante* or implicitly recognized *ex post*. Although conditionality can be a useful commitment device, it cannot substitute for country ownership. Conditionality should be used with judicious selectivity and tailored to country circumstances. Moreover, its limitations and opportunities must be recognized in designing a new generation of conditionality.

Because economic policies are driven primarily by domestic political processes, conditionality is appropriate only if there is commitment and capacity to reform. Research on aid effectiveness emphasizes the importance of focusing adjustment support on countries with good policies and institutions. Building capacity and providing advice based on analytical work are better ways of nurturing reforms in their early stages. Moreover, although country ownership of reforms is critical, it is difficult to assess. A country's track record is among the most robust indicators of its readiness and capacity to reform. Though there are exceptions, the history of policy-based lending is littered with inaccurately assessed windows of opportunity for reform.

Enhancing the development effectiveness of policy-based lending also requires making judicious use of the various design options that allow conditionality to be customized to country circumstances, including in terms of the number and nature of conditions and the phasing and tranching. Most attempts to address performance deficiencies and capacity limitations through a larger number of more complex conditions have been ineffective. Thus, conditionality should focus on priorities grounded in country ownership and capacity and be limited to policy and institutional actions under the control of the executive branch.

Although research on aid effectiveness encourages phasing and tranching based on performance rather than promises, a pure outcome focus is fraught with practical difficulties and is no substitute for conditionality. Good practice also suggests the need for a clear medium-term framework for policy-based lending linking policy actions, progress indicators, and expected outcomes. In addition, conditionality customized to country circumstances implies strengthening countries' capacity to monitor and evaluate progress toward development objectives.

These principles are embedded in the programmatic approach to policy-based lending, which recognizes the importance of country ownership through increased flexibility, reflects countries' track records, and supports sustained engagement for complex medium-term institutional and policy reforms. This approach to conditionality does not just involve providing financing, it also serves as a vehicle for policy dialogues that typically involve transfers of advice and knowledge. Programmatic policy-based lending calls for a prudent mix of government policies, intermediate benchmarks, and ultimate outcomes embedded in a medium-term policy framework and involving a series of operations linked by specific but flexible triggers. Thus it has emerged as a promising way to reconcile the debate between the traditional *ex ante* approach and the aspirations of the results-based approach to conditionality.

Notes

Stefan G. Koeberle is Advisor in the Operations Policy and Country Services (OPCS) Vice Presidency at the World Bank. The author's e-mail address is skoeberle@worldbank.org. This article draws on work undertaken as part of OPCS research on adjustment lending. The author gratefully acknowledges advice from Harold Bedoya, Bruce Fitzgerald, and Ulrich Zachau and research assistance from John Factora and Thaddeus Malesa.

1. World Bank Operations Policy and Country Services calculations based on OED data.

2. World Bank Operations Policy and Country Services calculations based on World Bank ALCID.

3. OED defines *demandingness* as the extent to which the project could be expected to strain the economic, institutional, and human resources of the government or implementing agency. *Complexity* refers to such factors as the range of policy and institutional improvements contemplated, the number of institutions involved, the number of project components and their geographic dispersion, and the number of cofinanciers. *Riskiness* refers to the likelihood that the project as designed would be expected to fail to meet relevant project objectives efficiently.

4. World Bank Operations Policy and Country Services calculations based on World Bank ALCID.

5. World Bank Operations Policy and Country Services calculations based on World Bank ALCID and OED data.

6. An important factor that allows borrowers to take on responsibility for monitoring reform implementation is to assign performance indicators to specific government agencies—ideally the same agencies responsible for implementing the program or its components. The importance of adequate monitoring is reflected in the proposed update of the Bank's operational policy on policy-based lending (World Bank 2002b).

References

- Alesina, A., and D. Dollar. 1998. "Who Gives Foreign Aid to Whom and Why?" NBER Working Paper 6612. National Bureau for Economic Research. Cambridge, Mass.
- Burnside, C., and David D. 1997. "Aid, Policies, and Growth." Policy Research Working Paper 1777. World Bank, Washington, D.C.
- Collier, P. 1997. "The Failure of Conditionality." In C. Gwin and J. Nelson, eds., *Perspectives on Aid and Development*. Policy Essay 22. Washington, D.C.: Overseas Development Council.
- Collier, P., P. Guillaumont, and S. Guillaumont. 1997. "Redesigning Conditionality." *World Development* 25(9):1399–1407.
- Devarajan, S., D. Dollar, and T. Holmgren. 2001. *Aid and Reform in Africa*. Washington, D.C.: World Bank. Available online at www.worldbank.org/research/aid/africa/intro.htm.
- Easterly, W. 2001. *The Elusive Quest for Growth*. Cambridge, Mass.: MIT Press.
- Emblad, S., and H. Gilles. 2000. "Conditionality Revisited: A New Approach in Burkina Faso." PREM-note 35. World Bank, Washington, D.C. Available online www1.worldbank.org/prem/PREMNotes/premnote35.pdf.
- European Commission. 2001. "Conditionality Reform: Some Results of the Burkina Faso Pilot Case." Presented at the OECD Seminar on Aid Effectiveness, Selectivity and Poor Performers, 17 January. Paris. Available online at www.oecd.org/pdf/M00021000/M00021252.pdf.
- Filmer, D., J. S. Hammer, and L. Pritchett. 2000. "Weak Links in the Chain: A Diagnosis of Health Policy in Poor Countries." *World Bank Research Observer* 15:199–224.
- Gunning, J. W. 2000. "The Reform of Aid." Oxford University, Centre for the Study of African Economies.

- Haggarty, L., and Y. Matuda. 1999. "Assessing Clients' Commitment to Sectoral Reforms: A Reform Readiness Analysis." World Bank, Development Economics Vice Presidency, Washington, D.C.
- Heaver, R., and A. Israel. 1986. "Country Commitment to Development Projects." World Bank Discussion Paper 4. Washington, D.C.
- IMF (International Monetary Fund). 2001a. "Structural Conditionality in Fund-Supported Programs." Washington, D.C. Available online at www.imf.org/external/np/pdr/cond/2001/eng/struct/index.htm.
- . 2001b. "Streamlining Structural Conditionality: Review of Initial Experience." Washington, D.C. Available online at www.imf.org/external/np/pdr/cond/2001/eng/collab/071001.pdf.
- . 2002. "Guidelines on Conditionality." SM/02/276, Revision I. Washington, D.C. Available online at www.imf.org/External/np/pdr/cond/2002/eng/guid/092302.htm.
- IMF, OECD (Organisation for Economic Co-operation and Development), United Nations, and World Bank Group. 2000. *A Better World for All: Progress towards the International Development Goals*. Washington, D.C. Available online at www.paris21.org.
- Johnson, J. H., and S. S. Wasty. 1993. "Borrower Ownership of Adjustment Programs and the Political Economy of Reform." World Bank Discussion Paper 199. Washington, D.C.
- Kanbur, R., and T. Sandler. 1999. "The Future of Development Assistance: Common Pools and International Public Goods." Policy Essay 25. Overseas Development Council. Washington, D.C.
- Meltzer, A., C. F. Bergsten, C. W. Calomiris, T. Campbell, E. J. Feulner, W. L. Hoskins, R. L. Huber, M. H. Johnson, J. I. Levinson, J. D. Sachs, and E. E. Torres. 2000. *Results and Recommendations of the International Financial Institution Advisory Commission*. Final report to the U.S. Congress and Department of Treasury. Washington, D.C. Available online at www.bicusa.org/usgovtoversight/meltzer.htm.
- Milanovic, B. 2003. "The Two Faces of Globalization: Against Globalization as We Know It." *World Development* 31(4):667–83.
- Morrow, D. 1999. "Assessing Borrower Ownership Using Reform Readiness Analysis." PREMnote 25. World Bank, Washington, D.C.
- Mosley, P., J. Harrigan, and J. Toye. 1991. *Aid and Power: The World Bank and Policy-based Lending, Volume 1—Analysis and Policy Proposals*. New York: Routledge.
- SAPRI (Structural Adjustment Participatory Review Initiative). 2002. "The Policy Roots of Economic Crisis and Poverty." Washington, D.C. Available online at www.saprin.org/SAPRIN_Findings.pdf.
- SPA (Special Programme of Assistance for Africa) Task Team on Contractual Relationships and Selectivity. 2000. "Conditionality Reform: The Burkina Faso Pilot Case." Brussels. Available online at www.oecd.org/pdf/M00021000/M00021255.pdf.
- Stiglitz, J. 1998. "Towards a New Paradigm for Development: Strategies, Policies, and Processes." Prebisch Lecture presented at the United Nations Conference on Trade and Development, 19 October. Geneva.
- Thomas, V., A. Chhibber, M. Dailami, and J. de Melo, eds. 1991. *Restructuring Economies in Distress: Policy Reform and the World Bank*. New York: Oxford University Press.
- Thomas, V., M. Dailami, A. Dhareshwar, D. Kaufmann, N. Kishor, R. López, and Y. Wang. 2000. *The Quality of Growth*. New York: Oxford University Press.
- Wood, A., and M. Lockwood. 1999. "The 'Perestroika of Aid'? New Perspectives on Conditionality." Bretton Woods Project. Washington, D.C. Available online at www.brettonwoodsproject.org/topic/governance/poa2.pdf.
- World Bank. 1980. *Structural Adjustment Lending*. R80-122, IDA/R80-83. May 9, 1980. Washington, D.C.
- . 1986. "Structural Adjustment Lending: A First Review of Experience." Operations Evaluation Report 6409. Washington, D.C.

- . 1990. *Report on Adjustment Lending: Policies for the Recovery of Growth*. R90-57 IDA/R90-49. Washington, D.C.
- . 1992a. *Efficient Implementation: Key to Development Impact—Report of the World Bank's Portfolio Management Task Force (Wapenhans Report)*. Washington, D.C.
- . 1992b. *Third Report on Adjustment Lending: Private and Public Resources for Growth*. R92-47, IDA/R92-29. Washington, D.C.
- . 1994. *Adjustment in Africa: Reforms, Results, and the Road Ahead*. A World Bank Policy Research Report. New York: Oxford University Press.
- . 1996. *Issues in Adjustment Lending*. SecM96-18. Washington, D.C.
- . 1998a. *Assessing Aid: What Works, What Doesn't and Why*. A World Bank Policy Research Report. New York: Oxford University Press.
- . 1998b. "Higher Impact Adjustment Lending (HIAL) in Sub-Saharan Africa: An Update." Africa Region, Chief Economist's Office, Washington, D.C.
- . 1999a. *Annual Review of Development Effectiveness: Toward a Comprehensive Development Strategy*. Operations Evaluation Department, Washington, D.C.
- . 1999b. "Higher Impact Adjustment Lending (HIAL)—Initial Evaluation." Operations Evaluation Report 19797. Washington, D.C.
- . 2000. *Report of the Working Group on Improving Quality of Monitoring and Evaluation in Bank Financial Operations*. CODE2000-72. Washington, D.C.
- . 2001a. *Adjustment from Within: Lessons from the Structural Adjustment Facility Review Initiative*. Washington, D.C. Available online at www.worldbank.org/research/sapri/WB_SAPRI_Report.pdf.
- . 2001b. *Adjustment Lending Retrospective*. Operations Policy and Country Services, Washington, D.C.
- . 2001c. *Supporting Country Development: Strengthening the World Bank Group's Support for Middle-Income Countries*. DC2001-0005. Washington, D.C. Available online at www.worldbank.org/html/extdr/pb/pbmiddleincome.htm.
- . 2001d. *World Bank Group Strategic Framework*. Washington, D.C.
- . 2002a. *Better Measuring, Monitoring and Managing for Development Results*. DC2002-0019. Development Committee, Washington, D.C.
- . 2002b. *From Adjustment Lending to Development Policy Support Lending*. Washington, D.C.
- . 2002c. "Supporting Country Development: World Bank Role and Instruments in Low- and Middle-Income Countries." DC/2000-19. Washington, D.C. Available online at [http://wbln0018.worldbank.org/DCS/DevCom.nsf/\(documentsattachmentsweb\)/September2000EnglishDC200019/\\$FILE/DC-2000-19-SCD.pdf](http://wbln0018.worldbank.org/DCS/DevCom.nsf/(documentsattachmentsweb)/September2000EnglishDC200019/$FILE/DC-2000-19-SCD.pdf).
- . 2003. "Rome Declaration on Harmonization: High Level Forum on Harmonization." Rome. Available online at http://siteresources.worldbank.org/NEWS/Resources/Harm-RomeDeclaration2_25.pdf.
- . Forthcoming. *Programmatic Adjustment Lending Retrospective*. Operations Policy and Country Services, Washington, D.C.
- World Bank and IMF. 1998. *Report of the Managing Director and the President on Bank-Fund Collaboration*. SecM98-733. Washington, D.C.
- . 2001. *Strengthening IMF-World Bank Collaboration on Country Programs and Conditionality*. Washington, D.C. August 2001.
- . 2002. *Strengthening IMF-World Bank Collaboration on Country Programs and Conditionality—Progress Report*. Washington, D.C.

World Bank Research Observer Cumulative Index, 2000–2003

Titles

- "A Comment on Cooper." Kenneth J. Arrow, 15, no. 2 (August 2000): 173–75.
- "Comment on 'Counting the World's Poor,'" Martin Ravallion, 16, no. 1 (Fall 2001): 149–56.
- "Comment on 'Counting the World's Poor,'" T. N. Srinivasan, 16, no. 1 (Fall 2001): 157–68.
- "Competition and Scope of Activities in Financial Services." Stijn Claessens and Daniela Klingebiel, 16, no. 1 (Spring 2001): 19–40.
- "Contagion: Understanding How It Spreads," Rudiger Dornbusch, Yun Chul Park, and Stijn Claessens, 15, no. 2 (August 2000): 177–97.
- "Corporate Performance in the East Asian Financial Crisis," Stijn Claessens, Simeon Djankov, and Lixin Colin Xu, 15, no. 1 (February 2000): 23–46.
- "Counting the World's Poor: Problems and Possible Solutions," Angus Deaton, 16, no. 1 (Fall 2001): 125–47.
- "Crisis, Adjustment, and Reform in Thailand's Industrial Firms," David Dollar and Mary Hallward-Driemeier, 15, no. 1 (February 2000): 1–22.
- "Deregulating the Transfer of Agricultural Technology: Lessons from Bangladesh, India, Turkey, and Zimbabwe," David Gisselquist, John Nash, and Carl Pray, 17, no. 2 (Fall 2002): 237–r65.
- "Developing Countries and a New Round of WTO Negotiations," Thomas W. Hertel, Bernard M. Hoekman, and Will Martin, 17, no. 1 (Spring 2002): 113–40.
- "Developing Countries' Changing Participation in World Trade," Will Martin, 18, no. 2 (Fall 2003): 187–203.
- "Ecology, History, and Development: A Perspective from Rural Southeast Asia," Yujiro Hayami, 16, no. 1 (Fall 2001): 169–98.
- "Evolution of Corporate Law and the Transplant Effect: Lessons from Six Countries," Katharina Pistor, Yoram Keinan, Jan Kleinheisterkamp, and Mark D. West, 18, no. 1 (Spring 2003): 89–112.
- "Financial Safety Nets: Lessons from Chile," Philip L. Brock, 15, no. 1 (February 2000): 69–84.
- "Foreign Bank Entry: Experience, Implications for Developing Economies, and Agenda for Further Research," George Clarke, Robert Cull, Maria Soledad Martinez Peria, and Susana M. Sánchez, 18, no. 1 (Spring 2003): 25–59.
- "The Gender Implications of Public Sector Downsizing: The Reform Program of Vietnam," Martín Rama, 17, no. 2 (Fall 2002): 167–89.
- "Globalization and the Labor Market," Martín Rama, 18, no. 2 (Fall 2003): 159–86.
- "Growth and Institutions: A Review of the Evidence," Janine Aron, 15, no. 1 (February 2000): 99–135.
- "How Effective Are Second-Generation Road Funds? A Preliminary Appraisal," Ken Gwilliam and Ajay Kumar, 18, no. 1 (Spring 2003): 113–28.
- "How Research Can Assist Policy: The Case of Economic Reforms in Uganda," John Mackinnon and Ritva Reinikka, 17, no. 2 (Fall 2002): 267–92.
- "The Impact of Financial Crises on Labor Markets, Household Incomes, and Poverty: A Review of Evidence," Peter R. Fallon and Robert E. B. Lucas, 17, no. 1 (Spring 2002): 21–45.

- "IMF Conditionality and Country Ownership of Adjustment Programs," Mohsin S. Khan and Sunil Sharma, 18, no. 2 (Fall 2003): 227–48.
- "Income Risk, Coping Strategies, and Safety Nets," Stefan Dercon, 17, no. 2 (Fall 2002): 141–66.
- "Industrial Policy: Growth Elixir or Poison?" Howard Pack, 15, no. 1 (February 2000): 47–67.
- "Infectious Diseases, Public Policy, and the Marriage of Economics and Epidemiology," Mark Gersovitz and Jeffrey S. Hammer, 18, no. 2 (Fall 2003): 129–57.
- "International Approaches to Global Climate Change," Richard N. Cooper, 15, no. 2 (August 2000): 145–72.
- "Judicial Reform," Juan Carlos Botero, Rafael La Porta, Florencio López-de-Silanes, Andrei Shleifer, and Alexander Volokh, 18, no. 1 (Spring 2003): 61–88.
- "Lessons in Structuring Derivatives Exchanges," George Tsetsekos and Panos Varangis, 15, no. 1 (February 2000): 85–98.
- "Paths of Institutional Development: A View from Economic History," Karla Hoff, 18, no. 2 (Fall 2003): 205–26.
- "The Politics of Russian Enterprise Reform: Insiders, Local Governments, and the Obstacles to Restructuring," Raj M. Desai and Itzhak Goldberg, 16, no. 1 (Fall 2001): 219–40.
- "Principles of Financial Regulation: A Dynamic Portfolio Approach," Joseph E. Stiglitz, 16, no. 1 (Spring 2001): 1–18.
- "Privatization and Corporate Governance: Principles, Evidence, and Future Challenges," Alexander Dyck, 16, no. 1 (Spring 2001): 59–84.
- "Privatization and Regulation of Transport Infrastructure in the 1990s," Antonio Estache, 16, no. 1 (Spring 2001): 85–107.
- "Productivity Growth and Sustainability in Post-Green Revolution Agriculture: The Case of the Indian and Pakistan Punjab," Rinku Murgai, Mubarak Ali, and Derek Byerlee, 16, no. 1 (Fall 2001): 199–218.
- "Public Intervention in Health Insurance Markets: Theory and Four Examples from Latin America," William Jack, 17, no. 1 (Spring 2002): 67–88.
- "Should Policy-Based Lending Still Involve Conditionality?" Stefan G. Koeberle, 18, no. 2 (Fall 2003): 249–73.
- "Social Capital: Implications for Development Theory, Research, and Policy," Michael Woolcock and Deepa Narayan, 15, no. 2 (August 2000): 225–49.
- "Social Investment Funds: An Organizational Approach to Improved Development Assistance," William Jack, 16, no. 1 (Spring 2001): 109–24.
- "Tight Money in a Post-Crisis Defense of the Exchange Rate: What Have We Learned?" Peter J. Montiel, 18, no. 1 (Spring 2003): 1–23.
- "Toward Transparency: New Approaches and Their Application to Financial Markets," Tara Vishwanath and Daniel Kaufmann, 16, no. 1 (Spring 2001): 41–58.
- "Trade, Foreign Direct Investment, and International Technology Transfer: A Survey," Kamal Saggi, 17, no. 2 (Fall 2002): 191–235.
- "Urbanization in Developing Countries," Vernon Henderson, 17, no. 1 (Spring 2002): 89–112.
- "The Washington Consensus a Decade Later: Ideology and the Art and Science of Policy Advice," T. N. Srinivasan, 15, no. 2 (August 2000): 265–70.
- "Weak Links in the Chain: A Diagnosis of Health Policy in Poor Countries," Deon Filmer, Jeffrey S. Hammer, and Lant H. Pritchett, 15, no. 2 (August 2000): 199–224.
- "Weak Links in the Chain II: A Prescription for Health Policy in Poor Countries," Deon Filmer, Jeffrey S. Hammer, and Lant H. Pritchett, 17, no. 1 (Spring 2002): 47–66.
- "Weathering the Storm: The Impact of the East Asian Crisis on Farm Households in Indonesia and Thailand," Fabrizio Bresciani, Gershon Feder, Daniel O. Gilligan, Hanan G. Jacoby, Tongroj Onchan, and Jaime Quizon, 17, no. 1 (Spring 2002): 1–20.
- "What Should the World Bank Think about the Washington Consensus?" John Williamson, 15, no. 2 (August 2000): 251–64.

Authors

- Aron, Janine. "Growth and Institutions: A Review of the Evidence." 15, no. 1 (February 2000): 99–135.
- Arrow, Kenneth J., "A Comment on Cooper," 15, no. 2 (August 2000): 173–75.
- Botero, Juan Carlos, Rafael La Porta, Florencio López-de-Silanes, Andrei Shleifer, and Alexander Volokh. "Judicial Reform," 18, no. 1 (Spring 2003): 61–88.
- Bresciani, Fabrizio, Gershon Feder, Daniel O. Gilligan, Hanan G. Jacoby, Tongroj Onchan, and Jaime Quizon. "Weathering the Storm: The Impact of the East Asian Crisis on Farm Households in Indonesia and Thailand," 17, no. 1 (Spring 2002): 1–20.
- Brock, Philip L., "Financial Safety Nets: Lessons from Chile," 15, no. 1 (February 2000): 69–84.
- Claessens, Stijn, and Daniela Klingebiel. "Competition and Scope of Activities in Financial Services," 16, no. 1 (Spring 2001): 19–40.
- Claessens, Stijn, Simeon Djankov, and Lixin Colin Xu. "Corporate Performance in the East Asian Financial Crisis," 15, no. 1 (February 2000): 23–46.
- Clarke, George, Robert Cull, Maria Soledad Martinez Peria, and Susana M. Sánchez, "Foreign Bank Entry: Experience, Implications for Developing Economies, and Agenda for Further Research," 18, no. 1 (Spring 2003): 25–59.
- Cooper, Richard N., "International Approaches to Global Climate Change," 15, no. 2 (August 2000): 145–72.
- Deaton, Angus, "Counting the World's Poor: Problems and Possible Solutions," 16, no. 1 (Fall 2001): 125–47.
- Dercon, Stefan, "Income Risk, Coping Strategies, and Safety Nets," 17, no. 2 (Fall 2002): 141–66.
- Desai, Raj M., and Itzhak Goldberg. "The Politics of Russian Enterprise Reform: Insiders, Local Governments, and the Obstacles to Restructuring," 16, no. 1 (Fall 2001): 219–40.
- Dollar, David, and Mary Hallward-Driemeier, "Crisis, Adjustment, and Reform in Thailand's Industrial Firms," 15, no. 1 (February 2000): 1–22.
- Dornbusch, Rudiger, Yun Chul Park, and Stijn Claessens, "Contagion: Understanding How It Spreads," 15, no. 2 (August 2000): 177–97.
- Dyck, Alexander, "Privatization and Corporate Governance: Principles, Evidence, and Future Challenges," 16, no. 1 (Spring 2001): 59–84.
- Estache, Antonio, "Privatization and Regulation of Transport Infrastructure in the 1990s," 16, no. 1 (Spring 2001): 85–107.
- Fallon, Peter R., and Robert E. B. Lucas, "The Impact of Financial Crises on Labor Markets, Household Incomes, and Poverty: A Review of Evidence," 17, no. 1 (Spring 2002): 21–45.
- Filmer, Deon, Jeffrey S. Hammer, and Lant H. Pritchett, "Weak Links in the Chain: A Diagnosis of Health Policy in Poor Countries," 15, no. 2 (August 2000): 199–224.
- Filmer, Deon, Jeffrey S. Hammer, and Lant H. Pritchett, "Weak Links in the Chain II: A Prescription for Health Policy in Poor Countries," 17, no. 1 (Spring 2002): 47–66.
- Gersovitz, Mark, and Jeffrey S. Hammer, "Infectious Diseases, Public Policy, and the Marriage of Economics and Epidemiology," 18, no. 2 (Fall 2003): 129–57.
- Gisselquist, David, John Nash, and Carl Pray. "Deregulating the Transfer of Agricultural Technology: Lessons from Bangladesh, India, Turkey, and Zimbabwe," 17, no. 2 (Fall 2002): 237–65.
- Gwilliam, Ken, and Ajay Kumar, "How Effective Are Second-Generation Road Funds? A Preliminary Appraisal," 18, no. 1 (Spring 2003): 113–28.
- Hayami, Yujiro, "Ecology, History, and Development: A Perspective from Rural Southeast Asia," 16, no. 1 (Fall 2001): 169–98.
- Henderson, Vernon, "Urbanization in Developing Countries," 17, no. 1 (Spring 2002): 89–112.
- Hertel, Thomas W., Bernard M. Hoekman, and Will Martin, "Developing Countries and a New Round of WTO Negotiations," 17, no. 1 (Spring 2002): 113–40.
- Hoff, Karla, "Paths of Institutional Development: A View from Economic History," 18, no. 2 (Fall 2003): 205–26.

- Jack, William, "Public Intervention in Health Insurance Markets: Theory and Four Examples from Latin America," 17, no. 1 (Spring 2002): 67–88.
- Jack, William, "Social Investment Funds: An Organizational Approach to Improved Development Assistance," 16, no. 1 (Spring 2001): 109–24.
- Khan, Mohsin S., and Sunil Sharma, "IMF Conditionality and Country Ownership of Adjustment Programs," 18, no. 2 (Fall 2003): 227–48.
- Koerberle, Stefan G., "Should Policy-Based Lending Still Involve Conditionality?" 18, no. 2 (Fall 2003): 249–73.
- Mackinnon, John, and Ritva Reinikka, "How Research Can Assist Policy: The Case of Economic Reforms in Uganda," 17, no. 2 (Fall 2002): 267–92.
- Martin, Will, "Developing Countries' Changing Participation in World Trade," 18, no. 2 (Fall 2003): 187–203.
- Montiel, Peter J., "Tight Money in a Post-Crisis Defense of the Exchange Rate: What Have We Learned?" 18, no. 1 (Spring 2003): 1–23.
- Murgai, Rinku, Mubarak Ali, and Derek Byerlee, "Productivity Growth and Sustainability in Post-Green Revolution Agriculture: The Case of the Indian and Pakistan Punjab," 16, no. 1 (Fall 2001): 199–218.
- Pack, Howard, "Industrial Policy: Growth Elixir or Poison?" 15, no. 1 (February 2000): 47–67.
- Pistor, Katharina, Yoram Keinan, Jan Kleinheisterkamp, and Mark D. West, "Evolution of Corporate Law and the Transplant Effect: Lessons from Six Countries," 18, no. 1 (Spring 2003): 89–112.
- Rama, Martín, "The Gender Implications of Public Sector Downsizing: The Reform Program of Vietnam," 17, no. 2 (Fall 2002): 167–89.
- Rama, Martín, "Globalization and the Labor Market," 18, no. 2 (Fall 2003): 159–186.
- Ravallion, Martin, "Comment on 'Counting the World's Poor,'" 16, no. 1 (Fall 2001): 149–56.
- Saggi, Kamal, "Trade, Foreign Direct Investment, and International Technology Transfer: A Survey," 17, no. 2 (Fall 2002): 191–235.
- Srinivasan, T. N., "Comment on 'Counting the World's Poor,'" 16, no. 1 (Fall 2001): 157–68.
- Srinivasan, T. N., "The Washington Consensus a Decade Later: Ideology and the Art and Science of Policy Advice," 15, no. 2 (August 2000): 265–70.
- Stiglitz, Joseph E., "Principles of Financial Regulation: A Dynamic Portfolio Approach," 16, no. 1 (Spring 2001): 1–18.
- Tsetsekos, George, and Panos Varangis, "Lessons in Structuring Derivatives Exchanges," 15, no. 1 (February 2000): 85–98.
- Vishwanath, Tara, and Daniel Kaufmann, "Toward Transparency: New Approaches and Their Application to Financial Markets," 16, no. 1 (Spring 2001): 41–58.
- Woolcock, Michael, and Deepa Narayan, "Social Capital: Implications for Development Theory, Research, and Policy," 15, no. 2 (August 2000): 225–49.
- Williamson, John, "What Should the World Bank Think about the Washington Consensus?" 15, no. 2 (August 2000): 251–64.