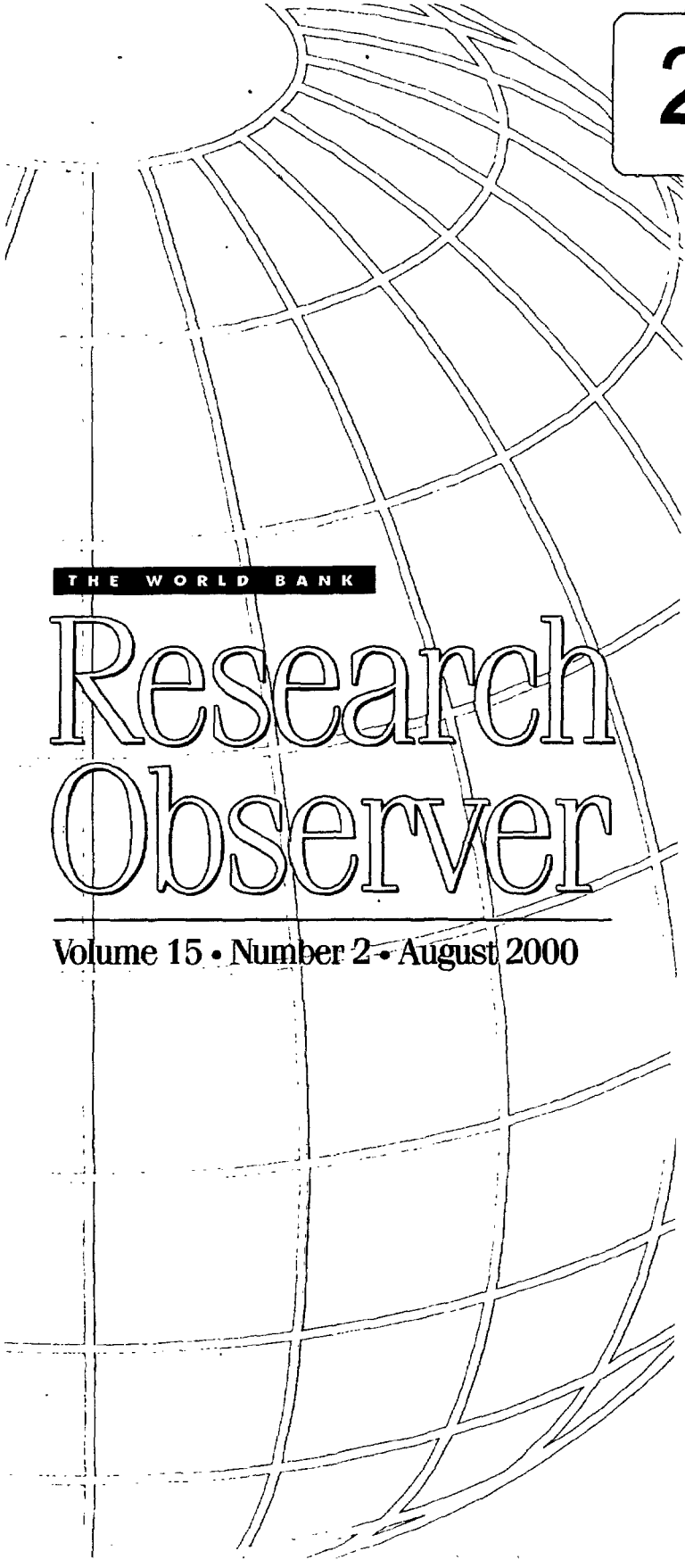


THE WORLD BANK

Research Observer

Volume 15 • Number 2 • August 2000

INSIDE

ISSN 0257-3032

THE WORLD BANK

Research Observer

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*, at the address in the copyright notice below.

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 the 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.

The World Bank Research Observer is published twice a year in February and August. Single copies may be purchased for US$12.95. Subscription rates are as follows:

	Individuals	Institutions
1-year subscription	US$25	US$40
2-year subscription	US$45	US$75
3-year subscription	US$60	US$100

Orders should be sent to: World Bank Publications, Box 7247–7956, Philadelphia, Pennsylvania 19170–7956, U.S.A. Readers with mailing addresses in developing countries and in socialist economies in transition are eligible to receive complimentary subscriptions on request. Written confirmation is required every two years to renew such subscriptions.

Manufactured in the United States of America
ISBN 0-8213-4610-5; ISSN 0257-3032

The World Bank Research Observer is indexed by the *Journal of Economic Literature*, the *Standard Periodical Directory*, ABI/INFORM, the Index to International Statistics, the Public Affairs Information Service, Inc., the Social Sciences Citation Index, and, online, by the Economic Literature Index and DIALOG. It is available in microform through University Microfilms, Inc., 300 North Zeeb Road, Ann Arbor, Michigan 48106, U.S.A.

THE WORLD BANK

Research Observer

Volume 15 • Number 2 • August 2000

International Approaches to Global Climate Change

Richard N. Cooper

This article surveys the issues involved in slowing the climate change induced by global emissions of greenhouse gases, especially carbon dioxide. It addresses the possible social and economic impacts of global warming, the elements involved in evaluating the pros and cons of steps to reduce those impacts, and the issues involved in engaging most of the world's states in a cooperative endeavor to reduce greenhouse gas emissions. It expresses doubts about the efficacy of a global approach based on national emission targets, such as those set by the 1997 Kyoto Protocol, and favors instead mutually agreed actions focused on a common emission tax. It also discusses issues of compliance with an international agreement to reduce emissions, actions states can take in the absence of international agreement, and contingency actions that might be considered if the problem proves to be more serious than now seems to be the case.

Human activity since 1800 has resulted in the emission of great volumes of gaseous materials into the atmosphere. Some of these gases—notably carbon dioxide, methane, and chlorofluorocarbons (CFCs)—absorb Earth's radiation, leading potentially to a warming of Earth's surface, which in turn could alter the world's climate. At the molecular level, CFCs are the most potent "greenhouse gases," but carbon dioxide has been emitted in greatest volume, largely from clearing forests and from burning coal and oil, and has the longest life in the atmosphere, thus accumulating over time. Atmospheric concentration of carbon dioxide in recent years has reached 360 parts per million (ppm), compared with about 280 ppm around 1800, and on some current projections is headed to 700 ppm (two and a half times preindustrial concentrations) by 2100 (Council of Economic Advisers 1998).

Wide scientific consensus suggests that under these conditions Earth's surface will become warmer on average, with temperature increases being higher in the higher latitudes. Sea levels will also rise, partly from melting glaciers but mainly from thermal expansion. Average global precipitation will increase.

The World Bank Research Observer, vol. 15, no. 2 (August 2000), pp. 145–72.

Beyond these general effects, the consensus dissipates. Earth's atmospheric physics and chemistry are complicated and not well understood; neither is the relationship between the atmosphere and the oceans nor that between climate and the biosphere (all forms of life). As a result, few agree on the rate at which chemical or biological processes take carbon dioxide out of the atmosphere, on the influence of greater warming and evaporation on cloud formation (which affects the extent to which the sun's rays are reflected away from Earth's surface), on the rate at which the oceans absorb heat from the atmosphere, and on a host of other relevant issues. Thus there is little agreement on either the ultimate extent or the pace of warming for any given trajectory of greenhouse gas emissions.[1]

Neither is the future trajectory of emissions known with high confidence, although continued "business-as-usual" economic growth can be expected to result in ever greater consumption of fossil fuels for many decades. The Intergovernmental Panel on Climate Change (IPCC 1996b) assumes in its main "business-as-usual" case a global growth of 1.1 percent a year in emissions of carbon dioxide between 1990 and 2100. At this rate, and in the absence of actions to reduce the growth of greenhouse gas emissions, the generally accepted range of warming over the course of the next century is 1.0–3.5°C, with a best guess of perhaps 2.0°C, and an equivalently uncertain rise in sea level of about half a meter.[2]

Confronted with these possibilities, the international community agreed at Rio de Janeiro in 1992 on a Framework Convention on Climate Change and at Kyoto in 1997 on a protocol that committed the countries listed in annex I of the framework convention to reduce emissions by the year 2012 to estimated 1990 levels or up to 8 percent lower, with the targets varying from country to country.[3] The countries listed in the framework convention include the 24 members of the Organisation for Economic Co-operation and Development (OECD) in 1995 plus Central and Eastern Europe and some successor states to the Soviet Union. The Kyoto Protocol has not yet come into force, and indeed the Clinton administration indicated it would not submit the document for Senate ratification until "significant participation" by leading developing countries was assured, as requested by Senate resolution.

Social and Economic Impacts

Several key potential economic effects of global warming have been identified. Perhaps the most important is the effect on global food production. But concerns have also been expressed about health, rising sea levels, and the general amenity of life. In addition, climate changes will have significant effects on nonhuman ecological communities.

Agriculture

The impact of climate on agriculture depends on the detailed effects of climate change, particularly the regional and seasonal changes in precipitation. But the adverse effects (if any) that might occur under that heading must be assessed against the facts that plants rely on carbon dioxide as a major input to production; that increased atmospheric carbon dioxide, taken alone, would actually increase agricultural yields; and that agricultural producers around the world, but especially in temperate zones, have a demonstrated capacity for adapting to a variety of changes in their economic as well as physical environment. Comprehensive work on agricultural response to global climate change is still in an early stage, but several studies suggest that changes of the likely magnitude will not have a significant impact on global food output, although significant effects might be experienced at the regional level. For example, Schimmelpfennig and others (1996) show that 20–30 percent declines in output of grains at two locations in the United States under substantial increases in temperature (4–5°C) are greatly moderated, or even converted into increases, with plausible adaptations by farmers; the carbon dioxide fertilization effect would ensure increases for all three products—maize, soybeans, and winter wheat (Reilly 1998:246). Similarly, Darwin and others (1995) show that under four climate models, declines in global production of cereals caused by a doubling of atmospheric carbon dioxide become modest increases when farmer adaptation and market adjustments are allowed. As table 1 shows, small declines in nongrain food production are more than offset by increases in cereal and livestock (Meyer and others 1998).[4]

Although global food production does not seem to decline with global warming—on the contrary—the regional distribution is uneven. Production rises in the higher latitudes, partly because of an increase in arable land, and tends to fall in the

Table 1. *Percentage Changes in Global Production of Cereals by Climate Change Scenario*

World	*No adaptation*	*Land use fixed*	*No restrictions on use*
GFDL	–23.5	–0.6	0.3
GISS	–22.6	0.2	0.9
OSU	–18.6	–0.5	0.2
UKMO	–29.3	–0.2	1.2

Note: "No adaptation" represents the additional quantities firms would be willing to sell at 1990 prices under the alternative climate. The next two columns represent changes in equilibrium quantities, under new equilibrium prices, with land use fixed and with no restrictions on land use. The results are based on equilibrium scenarios in four climate models that assume a doubling of atmospheric carbon dioxide: those developed at the Geophysical Fluids Dynamics Laboratory (GFDL), the Goddard Institute of Space Studies (GISS), Oregon State University (OSU), and the United Kingdom Meteorological Office (UKMO).

Source: Adapted from Darwin and others (1995); Reilly (1998).

Tropics, primarily because of an assumed decline in the availability of water. But the uncertainties must again be emphasized, particularly regarding regional effects, where the global climate models vary substantially in their projections.

Disease

Global warming may increase the threat of contagious diseases, which tend to thrive in warm climates. In particular, the potential range of endemic malaria will be extended as the relevant insect vector is able to move farther from the equator. But again, humans are not simply going to accept the increased spread of disease without strong reaction. Much more medical and pharmacological research is now devoted to temperate diseases and health conditions than to tropical diseases, largely because today's rich countries are mainly in temperate latitudes and they understandably pay most attention to the health conditions that most concern their residents. If malaria or other tropical diseases were to extend into these latitudes, one can assume that more resources would be devoted both to stopping the spread of the diseases and to immunizing the population against them. Advances in genetic engineering give added confidence that most diseases can be overcome or at least kept under control.

Moreover, the world economy will continue to grow; indeed, that is a key assumption underlying the projections of carbon dioxide emissions. Even a modest growth of 1 percent a year in global per capita income will result in a 170 percent increase in incomes over a century; a more likely growth rate of 1.5 percent would increase global per capita income by a factor of 4.4, with even more rapid growth in many regions that are now relatively poor. Increases in income enlarge the possible and likely human reactions to all aspects of the environment, including threats from disease. Malaria is virtually unknown in Singapore, near the equator, while it is common across the border in Malaysia. Greater wealth improves the capacity of both individuals and societies to control their environment.

Coastal Inundation

A rise in sea level will of course affect the habitability of coastal areas, where much of the world's population lives. A half-meter rise in sea level may not appear like much, but allowing for storm surges means that it could make currently inhabited areas uncomfortable, or even uninhabitable in extreme cases. Nicholls and others (1995) estimate that 5 percent of the world population would be affected, with 1 percent seriously at risk. But again, people will try to protect themselves by some combination of moving away and adjusting their structures and behavior. Nicholls and others estimate that a combination of such measures will reduce the population at risk by 88 percent, to 0.14 percent of the world population. Adaptation measures are

estimated to cost 0.056 percent of gross world product annually, or little more than 1/20th of 1 percent, although of course costs vary by region.

Market Impacts and Amenities

Various attempts have been made to assess the overall costs of climate change, and controversy surrounds the process. Among the most controversial issues are objections by some noneconomists to the insistence of economists on valuing the expected changes at market prices or something approximating them and the disagreements among economists on how best to do this. This is not the place to review the extensive and somewhat confused literature, but I do note that human behavior guided by foresight, or even by expectations based on one or two unpleasant experiences, can do a great deal to mitigate the costs of global climate change. To assume that people remain both ignorant and passive in the face of change is absurd; the entire international process involving the IPCC, the framework convention, and the Kyoto Protocol demonstrates that people are capable of thinking ahead and acting in anticipation—although not always wisely!

The emerging literature suggests that the best-guess costs associated with global warming are likely to be low, not catastrophic, as popular treatment of the subject sometimes suggests (table 2). Moore (1998), who assesses both the measurable and the not-so-easily measurable gains and losses for the United States, concludes that the United States is likely to be a net beneficiary of climate change—a result that would be even clearer for more northerly countries such as Canada and Russia. He argues that health is likely to improve in a warmer climate, that daily life would be more pleasant, and that recreational possibilities, although changed, would be changed in ways that cater more to current revealed preferences for recreation. Mendelsohn and Neumann (1999) have reached a similar conclusion.

The possibility that some countries may actually gain from climate change potentially complicates the prospects of a global agreement to limit the factors that lead to

Table 2. *Market Impacts with a 2.5°C Warming*
(percentage of gross domestic product)

Region	*Fankhauser*[a]	*Mendelsohn*[a]	*Tol*[a]
OECD	0.77	–0.17	0.27
Non-OECD	0.67	0.03	0.76
World	0.72	–0.18	0.52

Note: A minus sign signifies a net gain.

a. Mendelsohn assumes a 2.5°C rise in global mean temperature in 2060, whereas Fankhauser and Tol assume that this rise occurs in 2050. Note that only in Tol does damage depend on the rate of climate change. In all three cases, vulnerability is assumed as in 1990.

Source: Tol (1998).

climate change, since it will be difficult to persuade people to undertake costly or painful actions for benefits that accrue mainly to others.

Nonhuman Ecological Systems

Although human beings have demonstrated a remarkable capacity for adaptation to a variety of conditions and new developments, the same cannot be said about other species. Rapid climate changes may find many species unable to adapt in the time required.

Rising sea levels will affect natural ecosystems, particularly wetlands, which are known for their high levels of biological activity. Moreover, some human adaptation may come at the further expense of wetlands. For example, Nicholls and others (1995) estimate that without countervailing measures, rising sea levels will adversely affect 56 percent of the world's wetlands, and this figure rises to 59 percent when allowance is made for measures to protect human settlement. Existing wetlands are in rapid decline, however, for reasons that have nothing to do with climate change. If wetlands are to be preserved, affirmative human action will have to be taken, whether or not climate change threatens them.

Warming of temperate and northern latitudes will alter the natural vegetation, which in turn will alter the natural fauna. But trees take a long time to grow, and species move in nature only as rapidly as seeds can be carried by wind or creatures into newly habitable territory. Humans can assist other species to adapt to the new conditions, however, provided the requisite knowledge is available and the issue is considered sufficiently important.

Speciation is much higher in the Tropics than in higher latitudes. Microecosystems flourish in that climate, where highly specialized plants and (especially) animals with limited ranges have developed. Fortunately, temperature increases are likely to be least in tropical zones, but changes in precipitation and carbon dioxide fertilization will permit some species to flourish at the expense of others, possibly driving some to extinction.

A Framework for Collective Decisionmaking

Concerns about global climate change have led to pleas and indeed to some national commitments to slow or reverse the growth of greenhouse gas emissions. It is useful to identify the structural characteristics involved in attempting to mitigate global warming through formal collective action. There are three key features.

First, climate change brought about through an increased atmospheric concentration of greenhouse gases is a global issue, since whatever their earthly origin, the gases are widely dispersed in the upper atmosphere. Effective restraint must there-

fore involve all (actual and prospective) major emitters of greenhouse gases. Rich industrial countries account for most of the emissions today, but the Soviet Union was a major contributor before its dissolution and economic collapse in 1991, and the area can be expected to become a major source with economic recovery. Rapidly growing developing countries will become major contributors within a time frame that is relevant for managing the issue. Thus while the same requirements need not be imposed on all countries from the beginning, the agreement needs to be structured so that all countries will eventually participate. By one estimate, for example, full implementation of the Kyoto Protocol and continuation at the prescribed lower emission levels by annex I countries would slow the increase in average global surface temperature in 2050 by only 0.05°C, from an increase of 1.4° to 1.35°.

Second, the rewards from restraints on greenhouse gas emissions will come in the (politically) distant future, while the costs will occur in the political present. Moreover, the rewards are highly uncertain. The residents of some of today's countries, such as Canada, Russia, and perhaps the United States, may even expect to benefit from moderate climate change. It will thus be difficult to persuade people that they should make sacrifices in their own living standards for the sake of uncertain gains to their grandchildren and great-grandchildren, and to the grandchildren of others, remote in distance. The wide distribution of expected but distant benefits in response to collective action today provides an incentive for every country to encourage all to act but then to avoid acting itself—the so-called free-rider problem.

Third, the pervasiveness of the sources of greenhouse gas emissions—notably fossil fuel use, rice cultivation, and cattle production—implies that restraint will involve changes in behavior by hundreds of millions—if not billions—of people, and not merely by 180 or fewer governments, as in the typical treaty. Thus the most important part of an effective regime to limit climate change involves not the relationships among states but the effective influence of governments on the behavior of their citizens.

No major legally binding regulatory treaty touches all of these characteristics to the same degree. Typically, treaties apply to the actions of either governments themselves or a relatively few firms in a relatively few countries, as in the cases of halting nuclear testing or limiting production of CFCs. The Convention on International Trade in Endangered Species perhaps comes closest in its comprehensiveness; it requires states to prohibit international trade in an agreed list of products. The Chemical Warfare Convention is extremely intrusive in its monitoring requirements but has not yet come into force. The various agreements for management of international fisheries require cooperation from hundreds of fishermen, but with a few exceptions, these pacts have not been notably successful.

These three structural factors make collective decisions regarding actions to mitigate global climate change exceptionally difficult. Serious mitigation necessarily involves major reductions in the actual and prospective consumption of energy based

on fossil fuels (especially coal-fired electricity generation and the use of oil products for heat and transport). Because such consumption is at the very heart of modern industrialized economies, the costs of mitigation are both the economic and the psychological adjustments that must be made to move away from current energy systems and from wet rice and cattle production, which, along with leaks from gas and oil refining and distribution systems, are the main sources of methane from human activities. Moreover, the likelihood that the distribution of costs and benefits will be highly uneven across nations further complicates the task of reaching international agreement.

It is natural for an economist to compare the overall benefits of any proposed change with the overall costs required to make the change. Many noneconomists reject cost-benefit analysis as an artifact of calculators who ignore or underrate basic human values. But this rejection is simply an intellectual mistake; everyone who urges a change in policy (or resists one) is at least implicitly comparing costs with benefits. The disagreement, rather, is about how best to measure the alleged benefits and costs of the proposed change. Thus when Krause, Bach, and Koomey (1992) argue that on no account should the average global temperature be allowed to rise more than 2.5°C (the upper limit of Earth's temperature in the past 2 million years), implying, according to their worst-plausible-case calculations, that no more than 300 billion tons of additional carbon can be emitted into the atmosphere, they are implicitly arguing that the benefits of severe mitigation action are infinitely great and warrant any finite cost to achieve them. The authors are expressing an extreme degree of aversion to environmental risk. Others may properly disagree with such extreme valuation.

Table 3 illustrates the range of marginal benefits per ton of carbon emissions avoided and the range of marginal costs of reducing emissions by a ton of carbon for two different targets: a return to 1990 emissions and a 20 percent reduction from 1990 emissions. All the estimates must be taken as illustrative, as the methodologies used are quite different and are incomplete. With this caveat, these estimates show that costs generally exceed benefits, especially for 20 percent reductions in emissions. The wide variation suggests that the methodology for estimating costs and benefits can stand considerable improvement. Not surprisingly, the costs of reducing emissions rise with the magnitude of the reductions. Kram (1998) and Stavins (1999) suggest that the increases can be very steep.

The Discount Rate

Cutting greenhouse gas emissions involves incurring costs long before the benefits are registered. A comparison of near-term costs with future benefits requires use of a discount rate (or stream of rates) to put both benefits and costs into present value. Much has been written about the appropriate choice of a discount rate and the prin-

Table 3. *Selected Estimates of Benefits and Costs for Global Marginal Abatement of Carbon Dioxide*
(U.S. dollars per ton of carbon)

Benefit study	*Marginal benefit*[a]	*Cost model*[b]	*Marginal cost*	
			Stabilization	*20 percent reduction*
Ayres and Walter	30–35	Jorgenson-Wilcoxen	20	50
Nordhaus	7	Edmonds-Reilly	70	160
Cline	8–154	Manne-Richels	110	240
Peck and Teisberg	12–14	Martin-Burniaux	80	170
Fankhauser	23	Rutherford	150	260
Maddison	8	Cohan-Scheraga	120	330

a. For most studies the marginal benefit increases over time. The estimates presented here correspond to the period 2001–10.

b. Cost estimates are from a study by the Energy Model Forum of Stanford University, which ran 14 different cost models using common assumptions and standardizing for the emission reduction scenarios shown above.

Source: IPCC 1996a: (tables 6.11, 9.4).

ciples that should undergird the choice. Theoretical and some practical economists have been fascinated by the Ramsey model of savings, which suggests that the optimum social rate of time preference (r) can be expressed by the simple equation $r = \tilde{n} + \grave{e}g$, where $\tilde{n}$ is the pure rate of time preference, $\grave{e}$ is the elasticity of marginal utility with respect to additional consumption, and g is the growth rate of per capita consumption (see Nordhaus 1994 or IPCC 1996a, ch. 4, for an explanation). Plausible numbers for these variables lead to discount rates ranging from 0.5 to 3 percent (IPCC 1996a; Cline 1998).

The underlying rationale for avoiding or mitigating climate change is to benefit future generations. Yet to undertake investments in the near future that yield, say, 2 percent over the next century does a great disservice to future generations compared with other investments that we have strong reason to believe yield much higher returns many years into the future, if not for an entire century. Surely, in the interest of future generations, society should prefer high-return investments to low ones.

There is evidence that returns to education in developing countries exceed 20 percent (Psacharopoulos 1985, 1994). Returns to college education for a male in the United States reportedly equal 13 percent (Council of Economic Advisers 1996). A study of more than 1,000 projects completed by the World Bank in the 1970s and 1980s yielded an average (prospective) return of 16 percent (Pöhl and Mihaljek 1989). The World Bank and the U.S. government have stated threshold returns of 10 percent for evaluating prospective investments (recently reduced to 7 percent by the U.S. government). The corporate sector of the U.S. economy, one that is relatively rich in capital by global standards, yields an average pretax real return well above 10 percent. For all these reasons, 10 percent seems to be a reasonable rate of discount. A

high discount rate, of course, gives less weight to benefits (and costs) in the distant future. But that implication alone is not sufficient reason to reject it.

Maurice Scott of Oxford University has suggested 4 percent (reported in Beckerman 1996), partly because that has been the real yield on low-risk government bonds in recent decades. But even if resources can be extracted from the public at 4 percent, they should be invested in those activities with high (social) return. Only after we exhaust 10 and 7 and 5 percent opportunities should we accept investments with prospective yields of only 4 percent. Otherwise we deprive either future generations or our own generation unnecessarily.[5]

Some observers object to citing data on observed rates of return on grounds that actual decisions made today and in the past were not made under ideal conditions but instead reflect a number of imperfections both in markets and in the processes for making collective decisions. It would take me too far afield to explore this contention in relevant detail. Let me just stipulate that the real world is messy and that actual decisions (and market outcomes) deviate from any given set of ideal standards. The same observation applies to actions to mitigate greenhouse gas emissions. A plausible argument must be made that allowance for the various imperfections will raise the after-the-fact returns to mitigation actions relative to the observed returns on other investments.

The debate over the choice of a discount rate can be interpreted as an effort to reduce or eliminate imperfections in collective decisionmaking on public expenditures in general. But if such imperfections are important, and if other public investments seem to leave future generations still better off, low discount rates should also apply to those higher-yield investments. Advocates who apply them only to climate change must do so either because they believe the political prospects are better for improving collective decisionmaking on global climate change than for other, higher-yield public investments or because they must prefer mitigation of greenhouse gas emissions on some unstated grounds, not captured in the usual reckoning of costs and benefits over time, and want to support such actions—or both. In any event, it would be useful and desirable to open these considerations explicitly to wider discussion.

This sounds like common sense. What can be the objection to it? One possibility is that while the return on investment A (education, say) exceeds that on investment B (mitigation) in the near term, the reverse is true in the long run because of a secular decline in returns to investment A. Normally one could switch to B investments as returns to A drop below those to B. Investing in B now would be preferable only if for some reason it would be too late to switch to B investments later, after returns to new A investments fell.

This type of configuration is theoretically possible, but a plausible case must be made for both parts before one can conclude that A should be rejected in favor of B at the outset. In the standard neoclassical economic model, the returns to capital are

assumed to fall steadily as the ratio of capital to labor (and other factors) rises. But in historical—as distinguished from analytical—time, technical change has constantly increased the returns to (new) capital, and there is no reason to believe that the process will stop during the next century. Thus if returns to class A are high now relative to B, they are likely to remain so.

Distributional Considerations

The IPCC (1996a, ch. 4) authors seem to reject the efficiency argument that is emphasized here, not on the foregoing grounds but on the basis of equity. In essence, they argue that one cannot ethically say that investment A is superior to investment B even if it yields higher total future benefits if those who experience losses as a result of the investment are not actually compensated (in the absence of a social welfare function that indicates the relative weights that should attach to winners and losers).

This is a valid theoretical point. But if taken literally and applied seriously, it is a prescription for total inaction, especially when time frames as long as 100 years are under consideration. First, there is no collectively agreed social welfare function, and no prospect of one at a global level, so we cannot generally weigh winners against losers, especially over so long a time. Second, we cannot possibly know distant future winners and losers from our actions today. Try, for example, to identify the winners and losers from completion of the U.S. transcontinental railroad in 1869, or to forecast the winners and losers 100 years hence from construction of the Three Gorges Dam in China. And third, even if we had the requisite knowledge about future winners and losers and our preferences among them, we cannot bind future generations to adhere to those preferred outcomes. If we make rules, future generations can unmake them. If we plant trees, they can cut them down. If we consume less coal, they can consume more—and may actually do so because it is more readily available to them. The one legacy that cannot be reversed (short of a collapse of civilization) is enhanced knowledge—both a deeper understanding of nature and improved technology.

We should be concerned above all with passing more knowledge and higher incomes to the next generation than we received from our parents and should allow its members to decide how to distribute them. They will do so in any case, regardless of what we think. This is not to suggest that we should be completely indifferent to distributional effects. Our actions will affect the initial distribution of the next generation, and collectively we may want to avoid certain actions on grounds that we do not like their distributional effects. But here I mean the direct consequences of our actions, on which it may be possible to get collective agreement to avoid imposing extreme losses on certain classes of people. We cannot carry this logic into the more distant future, however, for the reasons already mentioned: we cannot possibly know the future impact on people (for one thing, we do not even know where they will be), and we cannot commit future generations to our preferences even if we did.

In any case, it is rather odd to urge costly action now for the sake of poor people in the distant future when we are not willing to take very costly action now for the sake of reducing poverty today. We have actual evidence on the amounts we are willing to spend, individually and collectively, in the name of economic development of today's poor countries: about 0.3 percent of the gross domestic product (GDP) of the rich countries, and only a portion of that is devoted to reducing poverty as such. If we are really concerned about the impact of possible future climate change on poor people, we should take more active steps to reduce their current poverty. That would improve their capacity to adapt to such climate changes as may take place and to take mitigation actions themselves.

If there is a general disposition within the rich countries to help people in poor countries, the best way to do it is probably through education. Education has at least three advantages with respect to mitigation of climate change. First, the rate of return seems to be substantially higher, at least on the estimates that have been made so far. Second, it is harder for future generations to undo the redistribution favored by this generation, since educated parents are likely to want to see their children educated. Third, education increases the capacity of any society, and of individuals, to adapt to changing circumstances, including but not limited to changes in climate.

Risk Aversion

It is widely taken for granted, at least on big issues, that people dislike uncertainty; they have an aversion to risk and are willing to pay something to reduce risk. This is the attitude that underlies the willingness of individuals to take out fire or liability insurance, to mitigate the possible costs of uncertain and perhaps even improbable unfavorable events.

The uncertainties associated with mitigating global climate change and its attendant costs are at least as great—and probably greater—than the uncertainties associated with other investments that could be made today; given risk aversion, one might thus conclude that costly mitigation actions should not be undertaken. However, the payoff from mitigation actions now will be greatest if the magnitude of global climate change and the associated costs turn out to be high, even if that is judged to be a low probability. Of course, if the costs associated with global climate change are low, any investment in mitigation actions will have a low or negligible return. But such investment may still be worthwhile as insurance against an uncertain but possibly costly contingency.

How do these considerations influence the discount rate? The precise answer is not at all straightforward, unless the uncertainty itself is related in a particular way to the passage of time. Roughly speaking, however, where an uncertain outcome (the future payoff from mitigation actions) is negatively correlated with the overall economic prospects and where the uncertainty grows exponentially with time, some

deduction from the discount rate used to evaluate mitigation actions is warranted. How much? That depends in detail on the nature of the uncertainty, an issue that needs much greater discussion, and on the degree of aversion to risk. But presumably it was this sort of consideration that led U.S. policymakers in 1980 to stipulate a discount rate of only 7 percent for publicly financed energy-related projects, 3 percentage points lower than the general standard for government investments. Serious disturbances in the field of energy, unlike other areas, can lower gross national product by a multiple, so some component of the energy investment can be regarded as an insurance premium designed to attenuate the economic impact of large disturbances in the world oil market.

Nordhaus (1994) models the determinants of climate change and the effects on emissions, temperature increase, warming damage, world output, and so on, focusing on optimal mitigation policies. He recalculates the optimal mitigation policy, taking into account uncertainties in those factors that influence climate change. Not surprisingly, the optimal reduction in greenhouse gas emissions, and the carbon tax required to achieve it, are higher in the presence of these uncertainties than they would be with confident best-guess projections. Concretely, the optimal carbon tax during the 1990s under the uncertain conditions postulated by Nordhaus is $12 a ton, compared with less than $5 a ton on the best-guess projection. Of course, the optimal policy is likely to change in response to new knowledge.

What about the possibility of truly disastrous outcomes as a result of global warming? Although the scientific community does not put a high probability on any of them, three are sometimes mentioned: first, warming sufficient to release the extensive methane contained in the Arctic permafrost, leading to a strong and possibly rapid reinforcement of warming; second, warming sufficient to break up the Antarctic ice dam and raise the oceans several meters, rather than half a meter, higher; and third, glacial melting in Greenland of a volume and character sufficient to deflect southward the warm North Atlantic currents, paradoxically making Europe a much *colder* place.

These possibilities, however remote, raise the question of risk aversion and how much insurance societies are willing to buy against improbable but highly costly contingencies. There is no doubt that individuals vary greatly in their degree of risk aversion and that commercial insurance policies do only a modest job of bringing these diverse preferences into harmony at the margin. The market for differences in preference regarding risk is much less well developed than the market to take advantage of differences in time preference. Each society has its own mechanism, through the political process, for deciding and acting on the degree of collective risk aversion. But the mechanism for the world as a whole is much less well developed, being mediated through diplomatic conferences such as those at Rio in 1992 and Kyoto in 1997, followed by public debate and ratification.

The political process, while essential for making decisions on collective risk, contains some serious weaknesses. Most notably, the discussion is not conducive

to honesty and straightforwardness. Some risk-averse parties will exaggerate the risks to persuade those who are less risk-averse than themselves. Others will attempt to minimize the estimated costs of early action or suggest that they can be borne by nonvoters (for instance, corporations). And some will cite legitimate concerns about greenhouse warming to encourage society as a whole to adopt a certain "lifestyle," such as relying less on the automobile, a technology that has been liberating for many people. By the same token, those who expect to bear the costs of political decisions in response to concerns about climate change will tend to minimize the risks and exaggerate the costs of mitigation—or even deny that a problem exists. In short, we should be on guard against strong but misleading or exaggerated arguments by all sides.

One way to deal with a potentially important problem that is subject to profound uncertainties is to establish a framework for action with broad participation and institutional procedures for integrating new information into decisions, as Schmalensee (1998) has urged, but to avoid costly actions that may turn out to be mistaken.

International Burden-Sharing

Suppose that the international community decides that steps should be taken to reduce greenhouse gas emissions and that many or all nations should be involved. What might be included in such a treaty? One approach, reflected in the Kyoto Protocol, is to impose national targets on emissions, possibly permitting some of the allowed emissions to be transferred from one nation to another. A second approach, which has received less attention, would call for a set of actions that states would agree to undertake. In my view, mutually agreed actions have better prospects of mitigating emissions than national targets.

Setting National Targets

When quantitative targets are imposed within countries, they almost universally respect recent history, being set at levels roughly in proportion to recent use (for instance, oil refinery throughput or emissions of sulfur or harvests of halibut). Targets based on emissions in a fixed base year such as 1990, as at Kyoto, have a similar character. They in effect allocate property rights to the existing tenants, accepting the right of ownership by virtue of possession or use. Targets allocated on this basis will be completely unacceptable, however, to countries that are or expect to be industrializing rapidly and that anticipate a disproportionately rapid growth in demand for fossil fuels. They will argue that most of the existing stock of greenhouse gases was emitted by rich countries and that those countries should therefore bear a disproportionate responsibility for cutting back. Thus developing countries did not

commit themselves to reduce emissions at Rio or Kyoto; within Europe, Greece and Spain expressed similar reservations.

Some observers have suggested that simple distributive justice would require that emission rights be based on population. Such an allocation would favor heavily populated poor countries such as Bangladesh, China, India, Indonesia, and Nigeria. To be meaningful in limiting climate change, such allocations would require drastic cutbacks in emissions by rich countries, implying radical reductions in living conditions there if implemented quickly. Targets based on population would be insensitive to varying resource endowments (such as hydroelectric power) and to the fact that countries depend on vastly different fuel mixes and have different levels of fuel consumption.

Reductions in living standards could be mitigated, but not avoided, by the sale of unused emission rights from poor to rich countries (discussed further below). But the financial transfers involved if emission rights were based on population would be immense relative to foreign assistance today—far greater than is likely to be politically tolerable. If carbon emissions were to take a plausible value of $100 a ton, for instance, the typical American family of four would have to pay $2,200 a year to sustain its current (direct and indirect) average level of emissions of about 26 tons a year, 22 tons over its per capita allocation (roughly 6 billion tons of carbon emissions a year divided by a world population of roughly 6 billion people). Total U.S. transfers to the rest of the world would amount to $130 billion a year, more than 10 times current U.S. foreign aid expenditures. Moreover, the transfers in practice would be made to governments, despite the underlying moral rationale for basing targets on population, and many people would question the desirability of transferring large sums to governments whose responsiveness to the needs of their own citizens has been indifferent or worse.

A natural compromise has been suggested: base the national targets on GDP (or recent past emissions) initially and gradually convert them to population-based targets over, say, 25 years. Here, however, we encounter some unpleasant arithmetic with respect to population-based emission rights. In 1995 India's per capita income (on a purchasing power basis) was about 5.2 percent that in the United States. Suppose that per capita income in India grows at 5 percent a year over the next 25 years and per capita income in the United States grows at 1 percent a year (this is a plausible scenario, although in reality the gap in growth rates is not likely to be so wide). Under those assumptions, Indian per capita income 25 years from 1995 (in 2020) would still equal only 14 percent of per capita income in the United States, and per capita consumption of energy would be many times higher in the United States than in India. Thus, after 25 years, one of three possible scenarios would emerge: India would not be effectively constrained; the United States would be very tightly constrained; or (under tradable emission permits) there would be huge transfers from the United States to India. The sense of global community is not likely to be great

enough by 2020 to sustain such large transfers—it is not that great *within* the United States today—and in any case such large transfers, either to governments or directly to citizens, would probably not be desirable, as some of the highly oil-dependent countries have discovered.

Perhaps the most reasonable way to allocate emission rights and the obligation to reduce emissions would be to calculate a "business-as-usual" trajectory of emissions for each country on the basis of recent history, development prospects, and past experience with the evolution of greenhouse gas emissions in relation to economic development. Then each country could be charged with reducing emissions by a uniform percentage, chosen in relation to global reduction requirements, relative to the assigned trajectory. Of course, even if this principle of allocation of rights and responsibilities were accepted as reasonable, the debate would simply shift to the choice of trajectories for each country. Developing economies aspire to grow rapidly. The Republic of Korea and Taiwan, China, have demonstrated that growth of more than 8 percent a year for three decades is possible. Most developing countries will set their aims similarly high and insist on energy-consumption growth to support them. They will be reluctant to accept lower emission targets without assurance that the technology will be available to achieve their growth targets with the lower emissions. Who is to say they are wrong?

Implementing the Agreement

Once national targets have been established, they must be translated into conditions that induce firms and households to change their consumption patterns to avoid certain activities. For large firms—generators of electricity, say—that could perhaps be done by fiat simply by setting quantitative limits for each generating plant. But for most economic agents the only practical way to alter behavior is to create price disincentives; that is, to tax the activities that generate the emissions.

Every international agreement must address the question of compliance and the associated question of monitoring behavior to discover if it deviates from the treaty requirements. In principle, given the objective, all significant greenhouse gases should be covered. In practice, given the many actors involved and the many sources of emissions, such broad coverage would be impossible to monitor and police. For practical reasons, therefore, attention is usually focused on fossil fuel consumption (plus a few other concentrated emitting activities, such as cement production). Monitoring such consumption is more or less manageable, since most of it must pass through some relatively narrow choke points (gas pipelines, oil refineries, electricity generating stations). Most coal production can be monitored at the mine-head or on the barges and railroads that transport the coal.

But this still leaves out a lot of emissions. Only about half of the greenhouse gas emissions (measured by radiative forcing, which is what is relevant for climate change)

since 1850 have come from the burning of fossil fuels (IPCC 1996b). The rest have resulted from the burning of tropical forests, cultivation of livestock and rice, dumping of garbage, and losses from gas pipelines. Omitting these sources would misrepresent total emissions. The Kyoto Protocol covers 24 gases, including methane and nitrous oxide, in addition to carbon dioxide (Council of Economic Advisers 1998). Monitoring emissions of all these gases will be difficult—and probably impossible—if developing countries are covered by the requirement.

If the fossil fuel carbon emission targets for rich countries are so demanding, how are they to be met? Conceptually, there are four ways: more efficient conversion of fossil fuels to usable energy in existing plants; switching away from fuels that are high in carbon per unit of energy produced (basically, shifting from coal to natural gas); building new plants that use less carbon per unit of usable energy (for example, nuclear power plants); and reducing end-user demand for energy. Unfortunately, the scope is limited for further change at the easiest monitoring points. Obsolete generating plants can be replaced with more efficient or less carbon dependent ones, but replacement demand in the OECD countries will be modest during the next 20 years, and replacing generating plants before they become obsolete is extremely expensive. In developing countries the demand for electric power is rising rapidly, so most of the generating capacity that will be available in those countries in 2010 could in principle be designed to use technology with reduced carbon dependence.

The consequence is that most of the reduction in the rich countries must come from downstream, at or near the points of final demand, where the number of consumers is greatest. Quantitative rationing is neither desirable nor feasible in market economies, so the reductions must be achieved by some combination of price (dis)incentives, exhortation through publicity, and education on best practice. Many consumers are not aware of the ways they can conserve energy without making radical changes in lifestyle. But in any case the key to success is not at the intergovernmental treaty level but rather in the incentives each government can provide to its own citizens. A treaty merely provides a vehicle for rough "burden-sharing" across countries and some international discipline in pursuit of the targets.

Opportunities for reducing emissions in new electric generating plants and other new industrial facilities will be greater, and the marginal cost lower, in developing countries than in mature economies. That realization has led to an emphasis on "joint implementation," a procedure whereby agents in rich countries can obtain credit against national targets in their own countries for making emission-reducing investments in developing countries. The idea is attractive. But under the Kyoto Protocol, the developing countries do not have national targets. Therefore, avoiding reductions in emissions in rich countries by investing in poor countries by itself will not reduce global emissions, since much investment must and will be undertaken in developing countries anyway. Reducing global emissions can be accomplished only by establishing detailed criteria for "additionality" in emission-reducing new invest-

ments—that is, by establishing norms by country and by project for least-cost power generating or energy-using investments and then counting reductions in emissions relative to such norms. The norms themselves would be changed as technology advanced. Establishing such norms would be both complicated and controversial because it would necessarily involve both judgment and approximation.

Imposing a Carbon Tax

There is an important alternative to setting national emission targets: an international agreement on a set of actions calibrated to achieve the desired emissions. Because to accomplish their quantitative objectives governments must in any case create the appropriate behavior-altering incentives for their citizens, and because setting a national allocation of global emission rights for both rich and poor countries is likely to prove so contentious as to be impossible, it may be easier simply to agree on a common use of instruments. For problems such as reducing emissions, the favorite instrument of economists is to tax the offending activity. All countries would agree to impose a common carbon tax, which would increase the price of fossil fuels in proportion to their carbon content. Such a tax would have at least two major advantages. First, it would encourage reduction of emissions where that can be done at least cost. All emitters would have the same incentive to act, but only those who saved more in tax payments than it cost to reduce emissions would undertake reductions; others would simply pay the tax. A carbon tax would encourage users everywhere to switch to natural gas (which would benefit Iran, Russia, and other countries with large gas reserves) and, more important, would prompt consumers generally to conserve fossil fuels. Second, a carbon tax would generate funds for governments that have trouble finding sources of revenue that do not have negative effects on economic incentives to work, save, or undertake commercial risks.

A common carbon tax would be easy to monitor. Enforcement would be more difficult, but all large countries except Cuba and the Democratic Republic of Korea hold annual consultations with the International Monetary Fund (IMF) on their macroeconomic policies, including the overall level and composition of their tax revenues, and the IMF could provide reports to the agency that monitors the treaty governing greenhouse gas emissions. Such reports could, if necessary, be supplemented by international inspection of the major taxpayers (for instance electric utilities) and of the tax agencies of participating countries.

Such a regime would present a major problem to democratic countries, however, because taxation goes to the heart of parliamentary prerogative, and most democracies will not welcome taxation by international agreement. Moreover, even modest energy taxes are politically unpopular.

Two additional problems need to be mentioned, neither insuperable. The first is that countries tax energy (especially oil) differently, and some countries continue to

price both coal and oil well below world levels. Should a uniform tax be levied on an uneven initial condition? If existing pricing practices are taken to reflect national preferences for allocating resources, a case can be made that the new carbon tax should be uniform, regardless of the initial tax burden. Of course, national policies would have to be monitored to ensure that the effect of the new tax was not undermined by other changes in tax or subsidy policy. Alternatively, the treaty could simply require a minimum national tax on emissions from fossil fuels, allowing existing taxes to drop toward that minimum, as advocated by Nordhaus.

The second problem concerns the disposition of revenue. Available estimates suggest that to have a significant impact on emissions, the tax might have to be substantial and thus would generate a great deal of revenue. To whom should this revenue accrue? Oil-producing states will suggest that if oil is to be taxed, they should levy it and get the revenue—indeed, that is what the attempts by the Organization of Petroleum Exporting Counties to control oil prices amount to. Oil-consuming countries, however, would feel doubly aggrieved if they charged more for oil to discourage its consumption and yet did not get the tax revenue; they would insist that the tax be levied on consumption and that the revenues accrue to them, not least so that they could reduce other taxes to ensure their continued prosperity and growth. In practice, the latter view is likely to prevail.

There is, however, a third possible claimant for the revenue: the international community. The international community has accepted a number of collective obligations that are cumulatively expensive. The most apparent are caring for refugees and peacekeeping operations, each of which cost the United Nations about $1.3 billion in 1995. Special assessments are now made for these activities, and several countries, including Russia and the United States, are in arrears. The regular UN budget runs $1.2 billion a year. In addition, donor countries finance the United Nations Development Programme and the International Development Association (at about $5 billion a year) to provide economic assistance to the poorest developing countries. Under the Rio framework convention, cooperation by developing countries in reducing emissions is conditioned on new financial support from the rich countries. Some or all of these activities could be financed in part by the tax revenues levied in pursuit of reduced emissions; obviously the major emitters, currently the rich countries, would pay most of the tax. But as poor countries develop, their contribution would increase automatically.

Estimates from several global energy-environment models suggest that a uniform reduction in carbon emissions from a "business-as-usual" baseline for each country or region would require very different carbon tax rates if that were the policy instrument used to reduce emissions. A uniform tax rate across the regions studied would result in quite different reductions from the baseline—as one would expect from the observation that countries around the world use energy with very different degrees of efficiency. Table 4 reports the per-ton carbon tax (in 1990 U.S. dollars) that would

Table 4. *Per-Ton Carbon Tax Required by 2050 to Achieve a 2 Percent Annual Reduction in Carbon Emissions from Baseline*
(1990 U.S. dollars)

Region	*Edmonds/ Reilly*	*Manne/ Richels*	*Green*	*Carbon rights trade model*
United States	1,096	208	340	754
Other OECD members	734	208	299	365
Former Soviet Union	325	990	180	2,245
China	341	240	67	1,109
Rest of world	1,012	727	329	763

Source: Dean and Hoeller (1993:153).

be required in five regions in the year 2050 to reduce carbon emissions by 2 percent a year from the baseline trajectory. Since the baseline trajectories project an increase in energy-related carbon emissions from roughly 6 billion tons a year in 1990 to 11 billion–19 billion tons in 2050, the 2 percent a year reduction would leave emissions that ranged between 3.3 billion and 5.7 billion tons in 2050, less than the 1990 levels.

By 2050 the world price (in 1990 dollars) of oil, 56 percent carbon by weight, is assumed to be $50 a barrel, two and half times its price in 1997. More recent projections suggest lower oil prices in the future, resulting from substantial improvements in the technology for discovering and extracting oil. Lower oil prices would lead to greater carbon emissions and thus to the need for greater taxes to bring them down to a target level. The price of coal, 75 percent carbon by weight, is assumed to be $60 a ton, about twice the recent price at points of exportation. Thus a tax in 2050 of $208 per ton of carbon would represent a 31 percent tax on oil and a 260 percent tax on coal. The loss in GDP engendered by this emission reduction program ranges (across the studies) from 1.3 to 4.9 percent in 2050 for the United States, from 2.3 to 6.4 percent for the countries of the former Soviet Union, and from 2.1 to 5.1 percent for the rest of the world—today's developing countries, minus China (Dean and Hoeller 1993). These results must be regarded as merely exploratory rather than definitive, but even the low estimates suggest a substantial cost to bringing energy-related carbon dioxide emissions below 1990 levels. The revenue these taxes would raise is also substantial. For instance, a carbon tax of $208 a ton in the United States would raise nearly $300 billion in revenue, 1.8 percent of estimated GDP in 2050. A carbon tax of $329 a ton in the rest of the world would raise $610 billion in 2050, nearly 3.2 percent of rest-of-world GDP in that year.

Trading Emission Rights

A gain in efficiency of emission reduction similar to that which would be achieved by a uniform world carbon tax can be achieved by allocating national targets to the

major emitters of carbon dioxide and allowing them to purchase or sell emission permits. A world market would quickly develop in such tradable emission permits, with a uniform world price. An emitter that could reduce emissions at a cost lower than the permit price would have an incentive to do so and would then sell its unneeded permits into the world market. An emitter that could reduce emissions only at a cost above the permit price could save money by buying enough permits to cover its excess emissions. Table 4 suggests that there would be much scope and mutual gain from a global market in permits, since the estimated costs of reducing emissions vary greatly from region to region in all the models.

The U.S. government, which has estimated the gains from trading emission rights, projects that the marginal cost for meeting the Kyoto target of 93 percent of 1990 emissions by 2012 would be about $200 a ton of carbon (calculated from Council of Economic Advisers 1998). If emission permits were allocated and traded, this cost would be reduced by 72 percent, to $56 a ton. Americans would not meet the 7 percent reduction themselves but rather would buy permits, mainly from Russia, which would have an easier time meeting its Kyoto target of no change from 1990 because of the collapse of heavy industry after 1990 and because of the considerable scope that country has for improving energy efficiency. Adding some key developing countries such as China and India to the trading regime would reduce the price to an estimated $23 a ton, allowing Americans to buy more permits more cheaply. The developing countries, by the same token, would receive substantial payments for their surplus emission rights.

An effective market in emission rights requires that the trading parties have well-defined property rights. These parties could in principle be governments, but in most countries permits would have to be allocated to the relevant firms. The allocation of emission rights would be a nontrivial political issue because of the distributional implications. As noted above, the historical tendency is to allocate quotas on the basis of historical performance. But the emission rights would have substantial value, and there would be little social merit in allowing the grandchildren of today's emitters to continue to own the rights 50 years from now. The distortions over time from "grandfathering" the emission rights could be avoided by auctioning the rights to the emitters from the start, with the revenues accruing to governments. Some of the revenues could temporarily subsidize today's emitters, which would have the major burden of adjustment in the new regime.

Because of the need to establish clear property rights to emissions, such a regime could include only annex I countries under the Kyoto Protocol; other countries have no emission targets. The inclusion of key developing countries would require them to agree to (necessarily growing) national targets and to a mechanism for allocating national targets to the emitting firms. Such allocation would of course be a strong temptation to graft and corruption because the rights would have substantial value under current estimates.

A permit trading regime would, moreover, require careful monitoring and enforcement to ensure that parties that had sold emission rights in fact cut their emissions to the levels stipulated. Another issue would be the potentially large transfers of wealth from permit-buying countries to permit-selling countries, with the magnitude of the transfer depending not only on the price of the permits but also on the initial allocation of emission targets, a matter touched on earlier.

Of course, implementation of the Kyoto targets without extension or global trading would also have distributional implications across countries, brought about both through the relocation of high-energy-using activities (and the associated investment) to countries without targets and through changes in the terms of trade that would result from implementing the Kyoto regime and making the secondary adjustments to that implementation.

The Kyoto Protocol, if taken seriously, will be costly to implement, will have a sharp impact, and will be of little benefit to the climate. Signatory governments have not leveled with their citizens on the full implications of implementing the agreement. When those implications are known, the public is likely to balk; it is plausible that most annex I countries will not meet the ambitious targets of the Kyoto Protocol by 2012. A cynic might argue that politicians understand this fully and have undertaken a classic straddle to have their cake and eat it too: cater to the single-minded constituents demanding immediate action on climate change, while committing themselves to an international framework that is likely to prove unworkable—not exactly a tragedy, since it may be neither necessary nor desirable.

Of course, any regime that taxes fossil fuels (directly or indirectly), that excludes some countries from the control regime, and that permits (or prohibits) trades in emission permits will have substantial redistributive effects among countries. Energy-intensive industries will move to countries that are not covered by the regime, hence influencing trade and investment flows. Countries' terms of trade will shift, especially away from exporters of fossil fuels. And, if trading is permitted, sales of emission permits will redistribute income from one country to another.

Compliance

Inevitably, the question of international coordination brings up the issue of free riders or noncompliers. Since the signatories are sovereign states and there is no overarching disciplinarian, as there is (in principle) within countries, fines or economic sanctions are the possible means of forcing compliance.

It would be difficult to prohibit trade related to the emission of greenhouse gases without in effect prohibiting trade with the offending country, since carbon dioxide–producing energy is required for virtually all production. But prohibiting trade could impose costs not only on the offending countries but also on their trading

partners that might well exceed the likely costs of global climate change. The advantages of one international regime would be sacrificed for another. Even if the threat worked, in the sense that countries were induced to comply with the emission objectives and the threat therefore did not have to be exercised, its existence might induce some important countries—China comes to mind—to reduce their dependence on trade as a matter of policy to avoid the possible cost of sanctions in the future, and that too would represent a cost of compliance.

Chayes and Chayes (1995) have argued that treaty law need not rely predominantly on sanctions, either in theory or in fact, and that indeed in many cases sanctions are counterproductive. They conclude that most actual or apparent deviations from treaty provisions arise from ambiguity and indeterminacy of treaty language, from limitations on the capacity of governments to carry out their undertakings, or from major changes in circumstances from those prevailing at the time the treaty was ratified. Deliberate violation of treaties is rare, and when it occurs on important issues, major participants exert extreme pressure outside the treaty for resumption of compliance, as happened with the violation of the Non-Proliferation Treaty by Iraq and the threatened withdrawal from that pact by the Democratic Republic of Korea.

The key factors in ensuring compliance are a commitment to the treaty objectives plus a high degree of transparency in governmental actions. Many regulatory agreements have a potential free-rider problem; countries are more likely to adhere to the provisions if other governments are seen to be adhering to the provisions, so a regular system for monitoring and reporting on the activities and actions covered by the treaty is important.

These days, the very legitimacy of many governments arises from their responsibility for international relations and their integration into the community of nations. External and, increasingly, domestic pressure will usually keep governments from deliberately flouting internationally agreed behavior. The need to engage publics in the reduction of greenhouse gas emissions, however, raises the issue of whether governments have the capacity to carry out their international commitments. Taxes are easier to monitor than quantitative emission targets.

Incomplete Steps toward Mitigation

Understanding the processes of climate change and their social and economic impacts is a daunting task; the world's knowledge of ways to reduce emissions with minimal social disruption and to disseminate best practices remains highly imperfect. But knowledge can be advanced along a variety of fronts through the actions of individual countries, as well as through internationally sponsored research. Much of this activity is under way, especially research on improved energy efficiency and alternatives to fossil fuels, but governments should ensure that no promising idea is

languishing for lack of funding. Knowledge will not resolve some issues, however, especially those concerning intergenerational distributional decisions and collective aversion to risk. These inevitably must be resolved through public discussion and political negotiations.

Even in the absence of an effective international agreement, countries may sensibly take steps to reduce greenhouse gas emissions. Subsidies and tax advantages that encourage the consumption of fossil fuels, especially—but not only—in developing countries, can be removed. Countries can encourage more rapid diffusion of available best practice in energy use through schools and public awareness programs. Sometimes outdated regulations need to be changed; for example, allowable rates in public utility regulation are often based on investment in new generating capacity but not on investment in conservation of electricity. Governments can provide funding for research in socially desirable new technologies and can, on their own, impose higher taxes on fossil fuels, devoting the revenues to reduction of other behavior-distorting taxes. There are many reasons other than inhibiting global climate change for adopting some or all of these measures: reduction of air pollution, reduction of urban congestion, and enhancement of energy security (especially with respect to imported oil). Reduction of greenhouse gas emissions would be a bonus, although a conscious one.

Further, international lending institutions such as the World Bank and the regional development banks are in the business of financing infrastructure projects in developing countries. Such infrastructure includes electric power generation and distribution, transport systems, and other major power-using activities. Once infrastructure is built, society adjusts to its long-lasting effects. Thus careful attention should be paid now to the longer-run social and economic implications of these investments. The extent and character of waste emissions, including greenhouse gases, should inform these investments, with special attention given to available best practice even when it is not considered by the principal contract-bidding firms. It is important to seek viable alternatives to coal-fired electricity generating plants, including a reexamination of the suitability of modern nuclear power technology with respect to safety, cost, and waste disposal.

A sensitive issue arises when investments in low-emission or nonemitting technologies cost *more* than the least-cost technologies (taking into account initial investment, maintenance, and lifetime input and waste disposal requirements). Should the lending institutions nonetheless insist on the investment that is more friendly to the environment? If so, who should pay for the incremental cost? It seems reasonable to decline to finance infrastructure investments that are unnecessarily damaging to the environment and acceptable to ask the borrowing country to pay fully for the incremental cost of any environmental benefits that accrue directly to it, but it is reasonable to expect the international community to pay most or all of the incremental cost

(depending on the income level of the borrowing country) associated with greenhouse gas emissions in cases in which the benefits will accrue to the world as a whole.

Contingency Planning

Many adverse developments *could* occur as a result of global climate change. It is much more difficult—if not impossible—to forecast with confidence what *will* actually happen. Some analysts have projected benign effects from global warming and easy adaptation to the adverse effects—especially for those whose income gives them room to maneuver. Thus developing countries understandably give higher priority to economic development than to averting climate change if the latter in any way inhibits development.

The great uncertainty about impacts, the prospect of serious gainers as well as losers, the high apparent cost of near-term actions to reduce emissions, and the need for eventual participation by countries with substantially different initial circumstances and hence greatly different priorities—all these factors make early action to stop the growth of greenhouse gas emissions, much less to lower them, highly problematic. Suppose the best guesses about climate change turn out to be too optimistic; or suppose that the forecasts are accurate but the international community is unable to reach agreement on costly, effective mitigation actions; or suppose that agreement is reached but countries prove unable to implement it. What will the community of nations do if experience reveals that climate change is severe and adverse? This possibility suggests the need for contingency planning to supplement research to develop cheap, low-emitting sources of energy and to satisfy human wants with lower requirements for energy. Such contingency planning can take two broad paths.

The first concerns how best to adapt to more serious climate change. It means not only pushing ahead with both the basic science and applied research for genetic engineering in many areas, especially agriculture, but also finding potential substitutes for possible useful species that may be lost. That search could be supplemented by a systematic program for collecting, cataloguing, and storing genetic material, mainly but not exclusively from plants, in the form of seed banks and DNA.

The second concerns slowing global warming as rapidly as possible. One route involves sequestration and even withdrawal of greenhouse gases, mainly carbon dioxide, from the atmosphere on a scale at least equal to continuing emissions. That approach will involve good stack absorbers and storage depositories of carbon dioxide. But it also might involve mobilizing the biosphere. Rapidly growing trees could be planted on a massive scale (by using planes to drop seeds, for example), especially as climate change extends the areas that can support them. More unconventionally,

barren portions of the oceans could be fertilized with the requisite minerals (thought mainly to be iron) so that microscopic carbon-loving plants can thrive.

A different approach would involve reducing the incidence of sunlight on Earth's surface, for example by placing reflecting surfaces in space or by increasing the albedo by altering cloud formation or by placing particulates in the atmosphere (through jet engine exhaust or by using cannons or rockets). Other possibilities will no doubt emerge over time. It is premature to commit to any particular method for rapid mitigation. Some suggestions will be impossibly expensive, and others will have unacceptable side effects, but it is important to encourage imaginative work on possible emergency actions.

Notes

Richard N. Cooper is professor of international economics at Harvard University.

1. The principal tool of analysis for global climate change is a large computer model (General Circulation Model, or GCM), of which several are in use, that attempts to model Earth's past and future climate as a function of received radiation, the characteristics of the atmosphere (such as concentration of carbon dioxide), and the dynamics of climate formation.

2. Excellent summaries of the scientific consensus and uncertainties about global climate change can be found in IPCC (1996b) and in Wuebbles and Rosenberg (1998). More recent work suggests that increased ice in Antarctica induced by a warmer climate (due to higher precipitation there) would withdraw water from the oceans and thus reduce this projected rise in sea level, possibly by as much as 50 percent (Thompson and Pollard 1997).

3. The European Union is treated as a single unit, with the maximum target reduction of 8 percent below emission levels of 1990. The United States agreed to a reduction of 7 percent and Japan to 6 percent, and Russia agreed not to exceed its 1990 level.

4. Fischer and Rosenzweig (1996) also find that global warming will increase global food production by 2050, with carbon dioxide fertilization playing an especially important role.

5. The comment by Kenneth Arrow misses my point. Suppose, for the sake of argument, that *all* resources for additional investment are taken from consumption and none from investment, as Arrow suggests. The future is still best served if new investment is undertaken that has the highest rate of return; an allocational mistake is made otherwise. Hence, the appropriate discount rate for evaluating potential projects is the rate of return on existing (marginal) investments.

I am concerned with incremental decisions. Arrow's approach would lead to a much higher rate of *total* investment. Indeed, with enough potential projects yielding more than his preferred consumption discount rate, it could lead logically to *all* output being used for investment—an absurd outcome. Before it was reached, the consumption rate of discount would rise sharply, at least enough to keep consumption well above subsistence levels. That would imply higher consumption rates of discount than Arrow either calculates or observes. Another factor is relevant: public investments must be covered, directly or indirectly, by taxation. Taxation in rich countries is already around 40 percent of GDP (higher in Europe, lower in Japan and the United States). In pursuit of rational preferences, the voting public is not likely to support proposals that require much higher levels of public investment with relatively low rates of return, which ultimately will require higher levels of taxation out of only modestly higher income. This argument is weakened if the "investment" itself produces revenue, which would be the case with a tax on carbon emissions.

References

Beckerman, Wilfred. 1996. *Through Green Colored Glasses.* Washington, D.C.: Cato Institute.

Chayes, Abram, and Antonia Handler Chayes. 1995. *The New Sovereignty: Compliance with International Regulatory Agreements.* Cambridge, Mass.: Harvard University Press.

Cline, William R. 1998. "Comment on Lind and Schuler." In Nordhaus (1998: 97–103).

Council of Economic Advisers. 1996. *Economic Report of the President.* Washington, D.C.: Government Printing Office.

———. 1998. "The Kyoto Protocol and the President's Policies to Address Climate Change: Administration Economic Analysis." Washington, D.C. Processed.

Darwin, Roy, Marinos Tsigas, Jan Lewandrowski, and A. Raneses. 1995. "World Agriculture and Climate Change: Economic Adaptations." Agricultural Economics Report AER-703. U.S. Agriculture Department, Washington, D.C.

Dean, Andrew, and Peter Hoeller, eds. 1993. *The Costs of Cutting Carbon Emissions.* Paris: Organisation for Economic Co-operation and Development.

Fischer, Cynthia, and Günther Rosenzweig. 1996. "The Impacts of Climate Change, Carbon Dioxide, and Sulphur Deposition on Agricultural Supply and Trade: An Integrated Impact Assessment." In N. Nakicenovic, W. D. Nordhaus, R. Richels, and F. L. Toth, eds., *Climate Change: Integrating Science, Economics, and Policy.* Laxenburg, Austria: International Institute for Applied Systems Analysis.

Intergovernmental Panel on Climate Change (IPCC). 1996a. *Climate Change 1995: Economic and Social Dimensions of Climate Change. Contribution of Working Group III to the Second Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, U.K.: Cambridge University Press.

———. 1996b. *Climate Change 1995: The Science of Climate Change. Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, U.K.: Cambridge University Press.

Kram, Tom. 1998. "The Costs of Greenhouse Gas Abatement." In Nordhaus (1998).

Krause, Florentin, Wilfred Bach, and Jonathan Koomey. 1992. *Energy Policy in the Greenhouse.* New York: John Wiley.

Mendelsohn, Robert, and James E. Neumann, eds. 1999. *The Impact of Climate Change on the United States Economy.* New York: Cambridge University Press.

Meyer, William B., W. Neil Adjer, Katrina Brown, Dean Graetz, Peter Gleich, John F. Richards, and Antonio Maghalães. 1998. "Land and Water Use." In Rayner and Malone, eds. (1998, vol 2).

Moore, Thomas Gale. 1998. *Climate of Fear: Why We Shouldn't Worry about Global Warming.* Washington, D.C.: Cato Institute.

Nicholls, R. J., S. P. Leatherman, K. C. Dennis, and C. R. Volanté. 1995. "Impacts and Responses to Sea-Level Rise: Qualitative and Quantitative Assessments." In R. J. Nicholls and S. P. Leatherman, eds., "The Potential Impact of Accelerated Sea-Level Rise on Developing Countries." *Journal of Coastal Research* 14(special issue).

Nordhaus, William D. 1994. *Managing the Global Commons: The Economics of Climate Change.* Cambridge, Mass.: MIT Press.

Nordhaus, William, D., ed. 1998. *Economics and Policy Issues in Climate Change.* Washington, D.C.: Resources for the Future.

Pöhl, Gerhard, and Dubravko Mihaljek. 1989. "Project Evaluation in Practice." World Bank. Processed.

Psacharopoulos, George. 1985. "Returns to Education: A Further International Update and Implications." *Journal of Human Resources* 20(fall):583–604.

———. 1994. "Returns to Investment in Education: A Global Update." *World Development* 22:1325–43.

Rayner, Steve, and Elizabeth L. Malone, eds. 1998. *Human Choice and Climate Change.* 4 vols. Columbus, Ohio: Battelle Press.

Reilly, John. 1998. "Comment on 'Climate Change Damages.'" In Nordhaus (1998).

Schimmelpfennig, David, Jan Lewandrowski, John Reilly, Marinos Tsigas, and Ian Parry. 1996. "Agricultural Adaptation to Climate Change: Issues of Long-Run Sustainability." Agricultural Economics Report AER-740. U.S. Department of Agriculture, Economic Research Service, Washington, D.C.

Schmalensee, Richard. 1998. "Greenhouse Policy Architectures and Institutions." In Nordhaus (1998).

Stavins, Robert N. 1999. "The Costs of Carbon Sequestration: A Revealed-Preference Approach." *American Economic Review* 89(September):994–1009.

Thompson, Stanley L., and David Pollard. 1997. "Greenland and Antarctic Mass Balances for Present and Doubled Atmospheric CO_2 from the GENESIS Version-2 Global Climate Model." *Journal of Climate* 10(May):871–900.

Tol, Richard S. J. 1998. "Comment on 'Climate Change Damages.'" In Nordhaus (1998).

Wuebbles, Donald J., and Norman J, Rosenberg. 1998. "The Natural Science of Global Climate Change." In Rayner and Malone (1998, vol. 2).

A Comment on Cooper

Kenneth J. Arrow

Richard Cooper has written an excellent survey on the economic implications of climate change, stressing the possibilities and limits of international policy. I comment here only on one part of the analysis, Cooper's choice of discount rates. I differ sharply with the basis of the analysis and wish to call attention to the extensive literature, dating back more than 40 years, on the choice of discount rates for public investment, which Cooper has disregarded.

Cooper refers to and then dismisses abruptly what may be called the *consumption* viewpoint. Investment is a sacrifice of consumption, and therefore the rate of return on a new investment should be at least equal to the implicit rate of return on consumption. (In this note, as in Cooper, all rates of return are real, not nominal.) This idea is hardly new; it is Marshall's "price of waiting." Böhm-Bawerk famously gave three grounds for the existence of a positive interest rate. First, if consumption is growing over time, the marginal utility of consumption must be falling; therefore, a sacrifice of consumption today must be compensated for by a greater increase in consumption in the future. Second, future consumption is automatically less valuable than the same consumption today, even if their marginal utilities are equal. Third, an increased lag of production behind inputs leads to an increase in production. The first two grounds together define the consumption rate of interest, as expressed in the formula (due, I believe, to Ragnar Frisch) $r = \rho + \theta g$, where ρ is the pure rate of time preference (corresponding to Böhm-Bawerk's second ground), θ is the elasticity of the marginal utility of consumption, and g is the rate of growth of consumption. The third is represented in modern language by the marginal productivity of capital, F_K, where $F(K)$ is output as a function of capital (taking labor and natural resources as given). The condition for optimal allocation over time, and the outcome of an intertemporal competitive equilibrium, is $r = F_K$.

Clearly, at an optimum, the rate of return on capital equals the consumption rate of interest, so that it would make no difference which rate is used for discounting. In fact, these two numbers seem to be very unequal. While estimating r is not straightforward, most estimates of the rate of return on consumption are on the order of 3 or 4 percent; the rate of return on capital is usually estimated (as by the U.S. Office of

The World Bank Research Observer, vol. 15, no. 2 (August 2000), pp. 173–75.

Management and Budget) as 7–10 percent. Cooper also cites studies which argue that the rate of return on certain other public investments is 10–12 percent. He concludes that these rates represent alternative uses of capital and therefore that efficiency demands that the rates found in private investment or alternative public investments be used to evaluate climate change policies, in particular, abatement.

I believe this argument is fundamentally flawed, as pointed out some time ago by Eckstein (1957). There is no reason why the investment to prevent climate change must be drawn from other kinds of investment. It can be drawn from consumption, and the consumption rate of interest is the alternative value of those resources. The more general situation is that the resources are drawn from both consumption and investment and therefore that the rate of return for evaluating public investment projects should be a weighted average of the consumption rate and the rate of return on capital. Since consumption is much larger than investment, it is reasonable to assume that the appropriate hurdle rate should be closer to the consumption rate.

The matter is further complicated by the fact that investments, public or private, spin off returns which, in turn, induce further saving and therefore investment, a point argued by Arrow and Kurz (1970). A full discussion of these points is well beyond the scope of this note, but the underlying structure of the approach has been set forth by Bradford (1975), and a complete explanation with a way of making the procedures operational is shown in Lind (1982).

Cooper quotes Maurice Scott as pointing out that the rate of return on risk-free securities is on the order of 4 percent. This observation is completely consistent with the previous estimates of the consumption rate of return. Cooper refuses to consider this rate as relevant, on the grounds that the investment should be made at the highest rate available. Of course, the correct conclusion is that *all* public investments that yield more than the consumption rate of return should be made, not that one should be made rather than the other. If the financing of public investments does not displace private investment, the rate of return on the latter is irrelevant.

The question that arises is, why does the market permit the discrepancy between the two rates of return? Two reasons are found in the literature. One is simply the corporate income tax, as stressed by Eckstein. The other is risk. For many reasons, including asymmetric information and moral hazard, the rate of return on private investments reflects a risk premium. The rate of return on public investments that we have been discussing is a riskless rate; the adjustment for uncertainty should come in the measurement of benefits and costs, not in the rate of return used in benefit-cost analysis. For these two reasons, the observed rate of return in the private sector is not the correct one for assessing public investment projects.

A final note: when considering investments for the very long future, the discussion has a strongly ethical component. I would argue that ethical preferences are subject to the same rules of rationality as private preferences, so that the *form* of the argument is much the same, although some of the parameters, particularly the rate of

pure time preference, may have different values ρ. But I believe the results will be quantitatively very similar.

Notes

Kenneth J. Arrow is Joan Kenney Professor of Economics, Emeritus and Professor of Operations Research, Emeritus at Stanford University.

References

Arrow, Kenneth J., and Mordecai Kurz. 1970. *Public Investment, the Rate of Return, and Optimal Fiscal Policy.* Baltimore, Md.: Johns Hopkins University Press for Resources for the Future.

Bradford, David. 1975. "Constraints on Government Investment Opportunities and the Choice of Discount Rate." *American Economic Review* 65:887–95.

Eckstein, Otto. 1957. "Investment Criteria for Economic Development and the Theory of Intertemporal Welfare Economics." *Quarterly Journal of Economics* 71:56–85.

Lind, Robert C. 1982. "A Primer on the Major Issues Relating to the Discount Rate for Evaluating National Energy Options." In Robert C. Lind, ed., *Discounting for Time and Risk in Energy Policy.* Baltimore, Md.: Johns Hopkins University Press for Resources for the Future.

Contagion: Understanding How It Spreads

Rudiger Dornbusch • Yung Chul Park • Stijn Claessens

Much of the current debate on reforming the international financial architecture is aimed at reducing the risks of contagion—best defined as a significant increase in cross-market linkages after a shock to an individual country (or group of countries). This definition highlights the importance of other links through which shocks are normally transmitted, including trade and finance. During times of crisis, the ways in which shocks are transmitted do seem to differ, and these differences appear to be important. Empirical work has helped to identify the types of links and other macroeconomic conditions that can make a country vulnerable to contagion during crisis periods, although less is known about the importance of microeconomic considerations and institutional factors in propagating shocks. Empirical research has helped to identify those countries that are at risk of contagion as well as some, albeit quite general, policy interventions that can reduce risks.

The financial turbulence that hit many East Asian countries in 1997 and then spread to other parts of the world continued unabated in 1998. Russia defaulted on its debt as confidence in global financial markets weakened. The turmoil roiled capital markets in industrial countries, dramatically altering the (relative) pricing of many financial instruments, and spilled over into speculative hedge-fund bets, leaving Long-Term Capital Management, a large U.S. hedge fund, facing near bankruptcy. The crisis subsequently hit Brazil, creating uncertainty about the country's ability to roll over its public sector debt, and spread to other emerging markets in Latin America and elsewhere.

International capital markets, particularly those in emerging markets, appear volatile, on both the downside and the upside. In the mid-1990s aggregate private capital flows into five crisis-affected East Asian countries (Indonesia, the Republic of Korea, Malaysia, the Philippines, and Thailand) averaged more than $40 billion annually, reaching a peak of about $70 billion in 1996. In the second half of 1997, more than $100 billion in short-term bank loans was recalled from these same five

The World Bank Research Observer, vol. 15, no. 2 (August 2000), pp. 177–97.

countries, as currencies and stock markets there collapsed. Capital flows reversed themselves again in 1999, and stock markets rebounded sharply across the region as portfolio and other foreign investors channeled resources back, slowing the reform process in some countries. The turmoil triggered recessions in many developing countries, most notably in Latin America (Perry and Lederman 1998); altogether, two-fifths of the global economy sank into recession in 1999, with the sharpest declines in gross domestic product concentrated in the developing world.

Neither the exact causes of this volatility nor the best international financial architecture for guiding the movement of international capital is yet known. Yet reducing volatility and contagion has been an important stated objective of recent reforms. Fischer (1998), for example, notes two important reasons for revamping the international financial architecture and smoothing the global economy. First, the high degree of volatility of international capital flows to emerging markets and these markets' limited ability to deal with this volatility make the recipient country vulnerable to shocks and crises that are excessively large, frequent, and disruptive. Second, international capital markets appear to be highly susceptible to contagion. Thus proposals to reform the international financial architecture must be based on a thorough understanding of the causes and consequences of contagion.

Episodes of volatility in international capital markets had occurred before the Asian crisis; an example was the "tequila effect" that followed Mexico's December 1994 devaluation and mainly affected Latin American countries. At that time, the issue of financial contagion had not yet caught the attention of policymakers in either industrial or emerging-market countries (but see Kindleberger 1989). Since the East Asian crisis, however, policymakers and economists have engaged in considerable research to identify and analyze the causes of financial contagion.

Contagion is best defined as a significant increase in cross-market linkages after a shock to an individual country (or group of countries), as measured by the degree to which asset prices or financial flows move together across markets relative to this comovement in tranquil times. An increase in comovement need not reflect irrational behavior on the part of investors. When one country is hit by a shock, liquidity constraints can force investors to withdraw funds from other countries. Because many financial transactions are conducted by agents rather than by principals, incentive issues also play a role in triggering volatility. A decision to pull funds from several countries can also reflect coordination problems among investors and insufficient mechanisms at the international level for dealing with countries' liquidity problems. Distinguishing among these various forms of investor behavior is very difficult in practice.

Although it is hard to determine whether comovements are irrational or excessive, empirical work has been able to document patterns in the vulnerability of countries to volatility and to identify possible channels through which contagion is transmit-

ted. Trade links, regional patterns, and macroeconomic similarities make countries vulnerable to volatility. Volatility can be transmitted from a particular country to other countries through common creditors and through actions of investors operating in international financial centers. These regularities have helped to identify countries that are at risk of contagion. Less is known about the importance of microeconomic conditions and institutional factors (including the actions of specific financial agents) in propagating shocks.

Governments and the private sector, as well as international financial institutions, must take action to minimize and manage the risk of financial contagion. But the balance is unclear. Should individual countries bear the burden of improving their financial sectors and enhancing the transparency of data, or is there a need to reform the rules under which international investors operate? Does contagion always represent fundamental factors, or should countries simply have more access to liquidity support to withstand the pressures of contagion? For answers, we must first look at what is known about the causes and transmission of contagion.

Contagion and Its Causes

Contagion refers to the spread of market disturbances—mostly on the downside—from one country to the other, a process observed through comovements in exchange rates, stock prices, sovereign spreads, and capital flows. In this article, we focus on contagion in emerging economies. The causes of contagion can be divided conceptually into two categories (Masson 1998; Wolf 1999; Forbes and Rigobon 2000; Pritsker 2000). The first category emphasizes spillovers that result from the normal interdependence among market economies. This interdependence means that shocks, whether of a global or local nature, can be transmitted across countries because of their real and financial linkages. Calvo and Reinhart (1996) term this type of crisis propagation "fundamentals-based contagion." These forms of comovements would not normally constitute contagion, but if they occur during a period of crisis and their effect is adverse, they may be expressed as contagion. Most empirical work seeks to explain the degree of comovements and the mechanisms for transmitting them—for example, how and under what conditions a speculative attack on a single currency is spread to other currencies on the basis of various fundamental relationships.

The second category involves a financial crisis that is not linked to observed changes in macroeconomic or other fundamentals but is solely the result of the behavior of investors or other financial agents. Under this definition, contagion arises when a comovement occurs, even when there are no global shocks and interdependence and fundamentals are not factors. A crisis in one country may, for example, lead investors to withdraw their investments from many markets without taking account of differences in economic fundamentals. This type of contagion is

often said to be caused by "irrational" phenomena, such as financial panics, herd behavior, loss of confidence, and increased risk aversion. But because these phenomena can be individually rational and still lead to a crisis, it is helpful to discuss each category in detail.

Fundamental Causes

Fundamental causes of contagion include macroeconomic shocks that have repercussions on an international scale and local shocks transmitted through trade links, competitive devaluations, and financial links.

COMMON SHOCKS. Studies identify various global shocks that can trigger market adjustments in an international context. A common global cause, such as major economic shifts in industrial countries and changing commodity prices, can trigger crises in—or large capital inflows to—emerging markets. Changes in U.S. interest rates have been identified with movements in capital flows to Latin America (Calvo and Reinhart 1996; Chuhan, Claessens, and Mamingi 1998). The strengthening of the U.S. dollar against the yen in 1995–96 was an important factor in the export downturn in East Asia and the subsequent financial difficulties there (Corsetti, Pesenti, and Roubini 1998; Radelet and Sachs 1998a, 1998b). In general, a common shock can lead to comovement in asset prices or capital flows.

TRADE LINKS AND COMPETITIVE DEVALUATIONS. Local shocks, such as a crisis in one economy, can affect the economic fundamentals of other countries through trade links and currency devaluations. Any major trading partner of a country in which a financial crisis has induced a sharp currency depreciation could experience declining asset prices and large capital outflows or could become the target of a speculative attack as investors anticipate a decline in exports to the crisis country and hence a deterioration in the trade account.

Competitive devaluations can be another channel for transmitting contagion. Devaluation in a country hit by a crisis reduces the export competitiveness of the countries with which it competes in third markets, putting pressure on the currencies of other countries, especially when those currencies do not float freely. According to Corsetti and others (1999), a game of competitive devaluation can induce a sharper currency depreciation than that required by any initial deterioration in fundamentals. In addition, the noncooperative nature of the game can result in still greater depreciation compared with what could have been attained in a cooperative equilibrium. If market participants expect that a currency crisis will lead to a game of competitive devaluation, they will naturally sell their holdings of securities of other countries, curtail their lending, or refuse to roll over short-term loans to borrowers in those countries. This theory gains some credence from the fact that during the East

Asian crisis in 1997, exchange rates depreciated substantially even in economies such as Singapore and Taiwan, China, which did not necessarily appear vulnerable to a speculative attack on the basis of their fundamentals.[1]

FINANCIAL LINKS. Economic integration of an individual country into the world market typically involves both trade and financial links. Thus a financial crisis in one country can lead to direct financial effects, including reductions in trade credits, foreign direct investment, and other capital flows abroad. For example, firms in East Asia that are linked to, say, Thailand by trade, investment, and financial transactions would be adversely affected if a crisis were to limit the ability of Thai firms to invest abroad, extend credit, and so on. Thus a financial crisis in Thailand would rationally be reflected in other countries, leading, for example, to comovements in asset prices and capital flows.

Investors' Behavior

The spread of a crisis depends on the degree of financial market integration. If a country is closely integrated into global financial markets, or if the financial markets in a region are tightly integrated, asset prices and other economic variables will move in tandem. The higher the degree of integration, the more extensive could be the contagious effects of a common shock or a real shock to another country. Conversely, countries that are not financially integrated, because of capital controls or lack of access to international financing, are by definition immune to contagion. In this sense, financial markets facilitate the transmission of real or common shocks but do not cause them. The actions of investors that are ex ante individually rational as well as collectively rational, even though they lead to volatility and may require policy changes, should be grouped under fundamental causes.

It can be argued, however, that investors' behavior, whether rational or irrational, allows shocks to spill over from one country to the next. The literature differs on the scope of rational versus irrational investor behavior, both individually and collectively. It is useful to start with a classification of types of investor behavior (see also Pritsker 2000). First, investors can take actions that are ex ante individually rational but that lead to excessive comovements—excessive in the sense that they cannot be explained by real fundamentals.[2] Through this channel, which can broadly be called investors' practices, contagion is transmitted by the actions of investors outside the country, each of whom is behaving rationally. Conceptually, this type of investor behavior can be further sorted into problems of liquidity and incentives and problems of informational asymmetry and market coordination. Second, cases of multiple equilibrium, similar to those in models of commercial bank runs, can imply contagious behavior among investors. Third, changes in the international financial system, or in the rules of the game, can induce investors to alter their behavior after an initial crisis.

LIQUIDITY AND INCENTIVE PROBLEMS. One form of rational behavior by individuals relates to liquidity and other constraints on lenders or investors. For example, the sharp currency depreciation and the decline in equity prices in Thailand and other economies affected early in the East Asian crisis resulted in large capital losses for some international institutional investors. These losses may have induced investors to sell off securities in other emerging markets to raise cash in anticipation of a higher frequency of redemptions. Liquidity problems may also face commercial banks whose lending is concentrated in particular regions. Suppose there is a single common creditor country with a heavy regional exposure, such as Japan in East Asia or the United States in Latin America. If banks from the common creditor country experience a marked deterioration in the quality of their loans to one country, they may attempt to reduce the overall risk of their loan portfolios by reducing their exposure to other high-risk investments elsewhere, possibly including other emerging markets in the region.

The incentive structure for individual financial agents can also create a tendency to sell off several markets at the same time. For example, an initial crisis may induce investors to sell off their holdings in other emerging countries because of their tendency to maintain certain proportions of a country's or a region's stock in their portfolios. As a result, equity and other asset markets in a range of emerging economies would also lose value, and the currencies of these economies would depreciate significantly. Schinasi and Smith (2000), for example, demonstrate that the value-at-risk models used by many commercial banks explain why financial institutions and other investors may find it optimal to sell most high-risk assets when a shock affects one of those assets. Although this type of behavior is individually rational, it can lead to overall adverse outcomes.[3] Garber (1998) analyzes the possible unpleasant dynamics associated with the use of unregulated financial derivatives in weak institutional settings.

Countries whose financial assets are widely traded in global markets and whose domestic financial markets are more liquid may be more vulnerable to financial contagion (Kodres and Pritsker 1998; Calvo and Mendoza forthcoming). Further, because global diversification of financial portfolios involves the cross-market hedging of macroeconomic risks, countries in which asset returns exhibit a high degree of comovement with a crisis-affected country in tranquil periods will be more vulnerable to contagion (Kaminsky and Reinhart 1998b).

These liquidity constraints and incentive structures could be important for all types of investors dealing with emerging markets. But it is possible that particular institutional investors—open-end emerging-market mutual funds, hedge funds, and proprietary traders—are especially susceptible to this type of behavior. Leveraged investors, such as hedge funds and banks facing margin calls, are more likely to confront liquidity problems in the wake of a crisis and be forced to sell their asset holdings in other markets. Managers of open-end funds may also need to raise li-

quidity in anticipation of future redemptions by investors. Faced with these problems, both leveraged investors and open-end-fund managers are likely to keep those assets whose prices have already collapsed and whose secondary markets have become less liquid and sell other assets in the portfolio. By doing so, investors cause other asset prices to fall, and the original disturbance can spread across different financial instruments and markets. The financial turmoil in the fall of 1998, when spreads on U.S. corporations rose from a normal level of 100 basis points to almost 200 basis points, suggests that these types of spillovers need not be limited to emerging markets but can also affect a broad spectrum of markets and borrowers.

INFORMATION ASYMMETRIES AND COORDINATION PROBLEMS. Another cause of contagion relates to imperfect information and differences in investor expectations. In the absence of better information to the contrary, investors may believe that a financial crisis in one country could lead to similar crises in other countries. A crisis in one country may then induce an attack on the currencies of other countries in which conditions are similar. This type of behavior can reflect rational as well as irrational behavior. If a crisis reflects and reveals weak fundamentals, investors may rationally conclude that similarly situated countries are also likely to face such problems; such reasoning helps explain how crises become contagious. This channel presumes, of course, that investors are imperfectly informed about each country's true characteristics and thus make decisions on the basis of some known indicators, including those revealed in other countries, which may or may not reflect the true state of the subject country's vulnerabilities. The information investors use may include the actions of other investors, which brings us to the effects of informational asymmetries on investor behavior.

Investors often do not have a full picture of the condition of every country as it affects their return on investment. In part, this limitation reflects the cost of gathering and processing information. Calvo and Mendoza (forthcoming) show that in the presence of information asymmetries, the fixed costs involved in gathering and processing country-specific information could lead to herd behavior, even when investors are rational. In their model, financial investors can be divided into two groups: informed and uninformed. Given the fixed cost of gathering and processing information, most small investors simply cannot afford to collect and process country-specific information individually (see also Agenor and Aizenman 1998). Instead, uninformed investors may find it less costly and therefore advantageous to follow the investment patterns of informed investors. In making asset choices, uninformed investors may then take into account portfolio decisions made by better-informed investors because such decisions provide useful market information.

Both informed and uninformed investors may tend to seek new information from those investors who acted earlier to adjust their portfolios. Thus if informed investors move to pull out of a country, the information cascade will lead less-informed

investors to disregard their own information and follow the informed investors, thereby causing even larger capital outflows (Scharfstein and Stein 1990; Wermers 1995; Calvo and Mendoza forthcoming). The tendency to herd may increase as the number of countries in which investments can be placed grows and the range of investors widens, thus raising the fixed cost of gathering and processing country-specific information. Some authors therefore argue that an increase over time in herd behavior may not be irrational (Banerjee 1992; Bikhchandani, Hirshleifer, and Welch 1992; Shiller 1995) and is instead an outcome of optimal portfolio diversification that becomes more prevalent as securities markets grow (Calvo and Mendoza forthcoming).

Another explanation for the increase in herding over time is that as investors have become more diverse and as establishing a reputation has become relatively more costly, investors may find it less expensive to follow the herd. Because some investors, particularly fund managers, may be more concerned about maintaining a reputation that depends on the performance of their portfolios, relative to that of a given market portfolio, than about their absolute performance, the risks of cascading behavior may be particularly high among institutional investors (see Kim and Wei 1997 for foreign exchange trading). Thus an individual institutional investor may refrain from acting first, even if market developments favor a new portfolio, for fear of losing his or her reputation if the decision should prove to be wrong. To be on the safe side, individual investors may follow the herd. All these outcomes involve behavior that is individually rational (albeit constrained) but that nevertheless can cause financial volatility.

MULTIPLE EQUILIBRIUMS. A more general explanation of contagion based on investors' behavior involves changes in expectations that are self-fulfilling in financial markets subject to multiple equilibriums. In this framework, contagion occurs when a crisis in one emerging market causes another emerging-market economy to move or jump to a bad equilibrium, characterized by a devaluation, a drop in asset prices, capital outflows, or debt default. In Diamond and Dybvig's (1983) model of bank runs, it is rational for individual depositors to either hold funds in the bank or withdraw funds, depending on the actions of all other depositors. The equilibrium result can be a bad outcome, that is, a run on the bank, or a good outcome, in which depositors keep their money in the bank. In an economic crisis, the result analogous to a bank run would be a sudden withdrawal of funds from a country sparked by investors' fears that unless they act quickly they will be too late to claim the limited pool of foreign exchange reserves.

Some observers argue that contagion is a consequence of sudden shifts in market expectations and confidence. Formal analytical models of multiple equilibriums have been developed to explain recent experience in emerging markets (Gerlach and Smets 1995; Jeanne 1997; Masson 1998). Such models, of course, do not lend themselves easily to empirical tests because the move or jump can be triggered by many factors,

some of which may appear to be fundamental causes. Drazen (1999), for example, shows that political factors may have played a role in the contagion during the 1992–93 Exchange Rate Mechanism crisis. And, of course, such changes in equilibrium are not limited to emerging markets but can also play a role in volatility and contagion in domestic financial markets.

CHANGES IN THE RULES OF THE GAME. Finally, contagion may result if investors change their assessment of the rules under which international financial transactions occur. The Russian default in 1998, for example, increased concern that other countries might follow similar unilateral policies regarding the treatment of foreign private creditors or that international financial institutions might not bail such creditors out as expected. The discussion on the international financial architecture itself following the East Asian financial crisis may have caused changes in the way investors viewed the rules of the game and weighed the odds of official bailouts. This concern is often alleged to have caused the turbulence in 1998 in Brazil (see Calvo 1998; Park 1998; Dornbusch 1999). Other reasons could include concern about the supply of funds from international lenders of last resort. In late 1998, for example, the International Monetary Fund (IMF) found itself called on to rescue so many countries that economists wondered whether it would be able to deal with many more liquidity crises. Thus a liquidity crisis in one country could trigger a run on other countries out of fear that the last eligible country would be out of luck.

Empirical Evidence of Contagion

Empirical examination of the evidence on contagion has focused mainly on comovements in asset prices rather than on "excessive" comovements in capital flows or disturbances in real markets. We discuss tests under the following categories: correlation of asset prices; conditional probabilities of a currency crisis; changes in volatility; comovements of capital flows and rates of return; and other tests.

Correlation of Asset Prices

The asset price tests measure the correlation among different economies in interest rates, stock prices, and sovereign spreads (Forbes and Rigobon 1999 survey the recent literature). A marked increase in correlations is considered evidence of contagion. Most of these studies find evidence of large comovements in a variety of asset returns, although there is less agreement on whether such comovements increase in the wake of a crisis. Several studies suggest that the Mexican crisis in 1994 was contagious. Calvo and Reinhart (1996) find that the comovement of weekly returns on equities and Brady bonds in emerging markets in Asia and Latin America was higher

after the Mexican crisis than before. Frankel and Schmukler (1998) show evidence that the prices of country funds in Latin America and East Asia displayed greater comovement with Mexican country funds after the crisis than before. Valdés (1997) confirms that the movements of secondary-market debt prices and credit ratings show that the Mexican crisis was contagious in Latin America. Agenor, Aizenman, and Hoffmaister (1999) report that the Mexican crisis had a sizable effect on movements in domestic interest rate spreads (and output) in Argentina. Baig and Goldfajn (1998) show that the cross-country correlations among currencies and sovereign spreads in Indonesia, Korea, Malaysia, the Philippines, and Thailand increased significantly during the East Asian crisis (from July 1997 to May 1998) compared with other periods.

A marked increase in correlations among markets in different countries may, however, not be sufficient proof of contagion. If markets are historically cross-correlated, a sharp change in one market will naturally lead to changes in other markets, and correlations during crises could increase appreciably. Forbes and Rigobon (2000) show that as volatility increases following a crisis, an increase in correlation could simply be a continuation of strong transmission mechanisms that exist in more stable periods. They also show that an increase in correlations of asset prices may result when changes in economic fundamentals, risk perception, and preferences are correlated, without any additional contagion. Because of this endogeneity, estimation of correlations must control both for comovement in these variables during normal times and for the effects of fundamentals in order to be able to identify pure contagion.

In practice, it is impossible to adjust for the effects of increases in volatility and endogeneity (as well as omitted variables) without making some more restrictive assumptions. Some papers have done so. Forbes and Rigobon (1999) investigate the evidence of contagion during the 1987 U.S. stock market crash, the 1994 Mexican peso crisis, and the 1997 East Asian crisis using daily data for stock indexes of up to 28 industrial countries and emerging markets. They show that correlation coefficients across multicountry returns are not significantly higher during crises, if one properly corrects for the problems of endogenous variables, omitted variables, and changes in the variance of residuals. Arias, Hausmann, and Rigobon (1998) also find only limited evidence of contagion.

In a test on the Exchange Rate Mechanism crisis, however, Favero and Giavazzi (2000) estimate a structural model of the behavior of European interest rates and find evidence of contagion in interest rate residuals even after controlling for normal interdependence. Using an autoregressive model, and thus controlling to some degree for structural relationships, Park and Song (2000) show that the Southeast Asian crisis did not directly trigger the crisis in Korea but that its fallout to Taiwan played an important role in the Korean crisis (see also Connolly and Wang 1998 and Tan 1998 for comovements of stock prices in Asia; Doukas 1989 for sovereign spreads).

Conditional Probabilities

Another way to control for the role of fundamentals is to study conditional correlation or probabilities, rather than raw correlations, and thus use a narrower definition of contagion. The most commonly used methodology, introduced by Eichengreen, Rose, and Wyplosz (1996) and Sachs, Tornell, and Velasco (1996), examines whether the likelihood of crisis is higher in a given country when there is a crisis in one or several other countries. This literature builds on studies in single-country crisis prediction (see Dornbusch, Goldfjan, and Valdés 1995; Sachs, Tornell, and Velasco 1996). Berg and Pattillo (1999) review this literature, and Goldstein, Kaminsky, and Reinhart (2000) provide a more general exposition of early warning systems.

The research involves estimating the probability of a crisis conditional on information on the occurrence of crisis elsewhere, taking into account fundamentals or similarities. One advantage of this definition of contagion is that it readily allows for statistical tests of its existence. These tests can also try to investigate the channels through which contagion may occur, distinguishing, among others, trade and financial links. Eichengreen, Rose, and Wyplosz (1996), using a probit model and a panel of quarterly macroeconomic and political data covering 20 industrial economies from 1959 through 1993, show that the probability of a domestic currency crisis increases with a speculative attack on a currency elsewhere and that contagion is more likely to spread through trade linkages than through macroeconomic similarities. Using a similar methodology, De Gregorio and Valdés (2000) conduct an extensive test of spillovers of the 1982 debt crisis, the 1994 Mexican crisis, and the Asian crisis using indexes of exchange rate pressures over three- and twelve-month horizons, real exchange rate movements, and changes in credit ratings.[4] They find that the Mexican crisis was the least contagious, while the Asian crisis was as contagious as the 1982 crisis (note that their methodology does not allow them to determine whether spillovers represent normal comovements or contagion). Importantly, they find that both debt composition and exchange rate flexibility limit the extent of contagion, whereas capital controls do not appear to curb it.

Taking an even longer perspective, Bordo and Murshid (2000) examine the record of financial crises over the past 120 years and the evidence of contagion in several macroeconomic variables. They find that the core countries of the prewar and interwar gold standards (the United Kingdom and the United States) appear to be important in disseminating shocks to the rest of the world but that such patterns actually appear to be weaker during crises. In contrast, after 1973, Bordo and Murshid find that countries that are otherwise not correlated show considerable comovement in asset prices during crises. They also find, however, that the volatility in correlation coefficients can be quite high; they are therefore reluctant to interpret the increase in correlations during recent periods as evidence of contagion, especially in light of the

Forbes and Rigobon (2000) finding that such increases might be normal. On the whole, these tests find no solid evidence that contagion has been increasing over time.

Glick and Rose (1998) apply a similar approach to five episodes of currency crises and 161 countries and find that trade linkages are important in propagating a crisis. They argue that contagion tends to be regional rather than global because trade tends to be more intraregional than interregional (see also Diwan and Hoekman 1999). Kaminsky and Reinhart (1998a) find that in terms of conditional probabilities, information on a large share of crisis countries in the sample increases the ability to predict a crisis elsewhere, particularly on a regional level. Their study further supports evidence that contagion has been primarily a regional phenomenon (see also Calvo and Reinhart 1996; Frankel and Schmukler 1998; Kaminsky and Schmukler 1999).

The evidence on the trade channel as an explanation of the regional nature of contagion appears more relevant to Latin America than to East Asia. Kaminsky and Reinhart (1998a) find a high probability that a crisis will spread through third-party linkages among Latin American countries (Brazil, Colombia, Mexico, and Venezuela), while similar linkages are not significant in East Asia. Brazil, Colombia, Mexico, and Venezuela have the largest share of bilateral trade with the United States among Latin American countries. Baig and Goldfajn (1998) analyze the trade matrix of East Asian countries and find that trade linkages among those countries are weak. They argue that trade linkages were not important in the expansion of the crisis in East Asia in 1997. Alba and others (1999) investigate the effects of competitive devaluations and argue that these alone could not have explained the large depreciation of other regional currencies after the Thai devaluation.[5] In transition economies, Gelos and Sahay (2000) find that correlations in pressures on the exchange market can be explained by direct trade linkages but not by measures of other fundamentals. They also find that market reactions following the Russian crisis look very similar to those observed in other regions during turbulent times. Tests thus find strong evidence that contagion is related to trade links and has a regional character.

Kaminsky and Reinhart (1998b) find that the probability of contagion increases when the crisis is associated with the common creditor channel. Indonesia, Malaysia, and Thailand are heavily dependent on Japanese commercial bank lending; a crisis in one or two of these countries spread to all three. Similar results are found in Latin America, where the conditional probability of a crisis in one country when several others are in crisis is estimated to be as high as 78 percent. Latin American countries obtain a large portion of credit from U.S. commercial banks. Analogous effects appear for other types of investors. Using data on closed-end country funds, Frankel and Schmukler (1998) test whether adverse shocks from the Mexican crisis were transmitted directly, or indirectly through financial markets based in New York. They find that Wall Street spread

the Mexican crisis to East Asian countries but did not play a role in its transmission to other Latin American countries.

Volatility Spillover

Another approach estimates spillovers in volatility—that is, cross-market movements in asset prices. Edwards (1998) examines Mexico's interest rate increase in 1994 and finds strong evidence of contagion from Mexico to Argentina but not from Mexico to Chile. Park and Song (1999) test volatility spillover among foreign exchange markets during the crisis period and find that the effects of the crises in Indonesia and Thailand were transmitted to the Korean foreign exchange market but that the Korean crisis did not reinfect the two Southeast Asian countries. These studies did not control for fundamentals and thus did not distinguish between a pure contagion and one based on fundamentals.

Capital Flows

Capital flows can offer the best insight into the transmission of contagion, but few tests of their comovements have been conducted. Van Rijckeghem and Weder (2000) test the role of bank lending and the effect of a common lender by examining capital flows to 30 emerging markets. In the Mexican and Russian crises, they find that the degree to which countries obtained funds from common bank lenders was a fairly robust predictor of both disaggregated bank flows and the incidence of a currency crisis. Froot, O'Connell, and Seasholes (2000) study the behavior of portfolio flows into and out of 44 countries from 1994 through 1998. They find strong evidence that price increases encourage portfolio flows and that price declines lead to reduced flows. They also find that regional factors such as common creditors appear to be increasingly important over time, suggesting that the actions of institutional investors could be a channel for transmission of shocks.

In an analysis of portfolios of mutual funds, Kaminsky, Lyons, and Schmukler (forthcoming) find that emerging-markets funds exhibit positive momentum. That is, they systematically buy winners and sell losers in both crisis and noncrisis periods, with one difference: contemporaneous momentum (buying current winners and selling current losers) is stronger during crises, whereas lagged momentum (buying past winners and selling past losers) is stronger during noncrisis periods. Contemporaneous momentum was at its strongest point during the 1994 crisis in Mexico. Importantly, Kaminsky, Lyons, and Schmukler find that mutual fund managers use contagion strategies; that is, they sell assets from any country when crisis hits another—strong evidence that contagion is transferred through the actions of portfolio investors. Choe, Kho, and Stulz (1999) find that foreign portfolio investors did not add to volatility (see also Kim and Wei 1999; Stulz 1999).

Other Tests

Most empirical papers find that macroeconomic weaknesses can provoke contagion because they make a country vulnerable to a crisis. Similarities in macroeconomic weaknesses can also lead to crisis because these signals are considered sorting devices and thus may induce a shift in investors' expectations. Ahlumawia (2000) attempts to separate the two effects and finds that after controlling for the direct effect of weaknesses, macroeconomic similarities can play a proximate role in contagious currency crises by coordinating investor shifts. A study of the behavior of the local lending activities of domestic- and foreign-owned banks in Argentina and Mexico reveals that foreign-owned banks may have had a stabilizing influence on overall credit growth in the banking sector, potentially reducing both countries' vulnerability to crisis (Goldberg, Dages, and Kinney 2000). There have been few tests using structural models to explain the degree of spillovers in real and financial markets. One is the application of a full trade model for crisis-affected East Asian economies (Abeysinghe 2000). Although transmission through trade played an important role, Abeysinghe found that the immediate economic contractions were largely a result of direct shocks attributable to pure contagion.

Implications and Reform Options

The empirical findings show that fundamentals help predict spillovers and that trade links are important factors as well. Common creditor and other links through financial centers transmit volatility from one country to another at a particular point. This work thus helps to identify those countries that are at risk of a spillover of volatility. Less is known about the importance of microeconomic conditions and institutional factors—including actions of specific financial agents and the various channels that induce spillovers—in propagating shocks. As a result, it has been difficult to attribute the spillovers to contagion. Importantly, the degree of spillover does not appear to have increased over time, and there are many similarities in the empirical regularities across periods and countries.

These findings suggest that comovements are unavoidable and that fundamental factors are important. To reduce the risks of financial contagion, reforms will thus be necessary. Many of these are of a general nature, such as reductions in fiscal and current account deficits, better management of exchange rates, improvements in financial sector services, enhancement of data transparency, and the like. Many economists have proposed, and some have analyzed, specific policy reform options to deal with contagion. Stiglitz and Bhattacharya (2000) argue, for example, that disclosure requirements may not be needed because markets can and do provide optimal incentives for disclosure. They also argue that under certain circumstances, information

disclosure could exacerbate fluctuations in financial markets and precipitate a financial crisis. Bushee and Noe (1999), looking at U.S. equity markets, find that improved disclosure by firms increases the volatility of their stock prices because the seemingly reduced information asymmetry and increased liquidity of the market attract more transient investors. Here Furman and Stiglitz (1998) point to the fact that even countries such as Sweden, with good regulation and supervision and transparent financial markets, have had financial crises.

Many economists also agree that although improved standards (for data disclosure, regulation, supervision, and corporate governance) could have prevented the buildup of vulnerabilities and reduced the risk of currency crises, they are only a first step. Improved implementation and surveillance are necessary as well. For example, Hawkins and Turner (2000), who analyze the role of prudential and other standards for financial institutions, stress implementation issues and predict that many developing countries will continue to have difficulty complying with what are essentially industrial-country standards.

For these reasons, several observers have argued for the use of prudential controls, particularly for financial institutions, to limit the risk of sudden capital outflows. Many countries already limit the maturity mismatches on foreign exchange liabilities and assets, monitor internal risk management systems of financial institutions, and issue sanctions for poor systems. Tightening could mean putting limits on the net open positions that financial institutions can take in foreign currency markets, as well as imposing limits on the amount of gross foreign currency liabilities (as a fraction of total liabilities or as a ratio to equity). Guidelines on internal risk management systems can be issued, and financial institutions can be more intensely monitored in this area. A further precautionary measure would require banks to hold more liquid foreign exchange assets relative to total foreign exchange liabilities than they are required to hold on domestic currency liabilities. And, finally, capital controls on (some type of) inflows at the country level might be useful to prevent the buildup of vulnerabilities; there is much less agreement in this area, however.

Specific reforms to the rules under which international investors operate are less apparent. There have been calls for limits on the operations of hedge funds, and revisions to the way in which commercial banks have to hold assets against short-term loans to emerging markets. But so far, no proposals specifically aimed at curbing the role of investors in contagion have emerged, let alone been agreed on. More discussion has occurred on the need to enhance liquidity support to withstand pressures of contagion, perhaps though an international lender of last resort or standstills on payments following a crisis. Clearly, whatever reforms are implemented, liquidity crises will still arise; thus a good part of the debate on the international financial architecture has focused on improving ways for dealing with the crises. In an analysis of the supply of international liquidity, Chang and Majnoni (2000) stress that liquidity provisions entail a tradeoff: liquidity provisions conditioned on certain poli-

cies and applied at penalty rates can deepen the possibility of a full crisis. At the same time, moral hazard concerns call for conditions and higher rates. Some new facilities—the Supplemental Reserves Facility, the Contingent Credit Lines of the IMF, the guarantee facility of the World Bank, and private sector facilities—are set up ex ante, which may reduce these concerns. They may also induce foreign investors to avoid generating a level of debt that may place the economy in a fragile situation.

Conclusion

Economists still do not know precisely what factors make countries vulnerable to contagion or the exact mechanisms through which it is transmitted at any given time. Although empirical evidence suggests that commercial banks and mutual funds can play a role, separating rational from irrational investor behavior is difficult in theory and in practice, as is determining whether irrational investor behavior is the sole source of contagion. Individually rational but collectively irrational behavior and (perceived) changes in the international financial system are likely to continue to have an influence. Further research—whether theoretical or empirical—on the role of international financial agents and the international financial system may shed light on these aspects. Such research could help identify characteristics that make countries vulnerable to contagion and could contribute to the development of specific policy prescriptions to reduce the risks of contagion, manage its impact, and help economies recover as efficiently as possible. In the meantime, it will be difficult to determine whether any measures—beyond strengthening the international financial architecture—can reduce the risks of contagion specifically.

Notes

Rudiger Dornbusch is professor of economics at the Massachusetts Institute of Technology, Yung Chul Park is professor of economics at Korea University, and Stijn Claessens is lead economist in the Financial Sector Policy Group of the World Bank. An earlier version of this paper was prepared for discussion at the World Institute for Development Economics Research workshop on financial contagion held at the World Bank, June 3–4, 1999, and reflects comments from participants.

1. An interesting question is whether Singapore and Taiwan, China, let their currencies depreciate to maintain export competitiveness or to conserve foreign reserves. Corsetti and others (1999) argue that these two economies were able to defend the original parities with their massive holdings of reserves and thus to withstand irrational withdrawal but were concerned about a loss of competitiveness. It can also be argued, however, that their decision to float their currencies was motivated by their efforts to fend off possible speculative attacks driven both by arbitrary shifts in expectations and by the reaction of panicky and irrational investors. Although the response may have been rational and optimal in either case, in that the perceived welfare costs of maintaining a stable exchange rate might have been too high, the contagion aspects and policy implications underlying the two rationales are quite different.

2. Investors can follow strategies that are ex ante irrational given their own preferences and the behavior of other investors. Although one cannot rule out the likelihood that this category is large, its lack of conceptual definition makes it difficult to analyze.

3. In a related argument, Goldfajn and Valdés (1997) find that when foreign investors withdraw deposits and loans, asset prices decline and asset markets become illiquid. Banks and other financial institutions thus risk failure because they cannot readily liquidate their assets. The liquidation problem may cause a run on these intermediaries themselves, provoking a banking or confidence crisis, and could lead to a speculative attack on the currency as foreign investors withdraw and convert their investments into foreign exchange. Such crises can spread to other countries when international investors are forced to sell off their positions in other national markets to make up for the liquidity shortage caused by the crisis in one country.

4. Caramazza, Ricci, and Salgado (1999) investigate the East Asian, Mexican, and Russian crises using an approach similar to that of Eichengreen, Rose, and Wyplosz (1996). They find that these crises do not differ much. Fundamentals, including trade spillovers, common creditors, and financial fragility, are highly significant in explaining crises, while exchange rate regimes and capital controls do not seem to matter.

5. In contrast, Baig and Goldfajn (1998) find large trade links among East Asian countries, which could explain some spillover based on reduced demand for intraregional exports (see also Huh and Kasa 1997).

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

"The Contagion Conference" refers to the conference "International Financial Contagion: How It Spreads and How It Can Be Stopped," sponsored by the World Bank, the International Monetary Fund, and the Asian Development Bank, February 3–4, 2000, Washington, D.C. Papers from the conference are available at <http://www.worldbank.org/research/interest/confs/past/papersfeb3-4/papers.htm>.

Abeysinghe, Tilak. 2000. "Thai Meltdown and Transmission of Recession within ASEAN4 and NIE4." The Contagion Conference.

Agenor, Pierre-Richard, and Joshua Aizenman. 1998. "Contagion and Volatility with Imperfect Credit Markets." *IMF Staff Papers* 45(June):207–35.

Agenor, Pierre-Richard, Joshua Aizenman, and A. Hoffmaister. 1999. "Contagion, Bank Lending Spreads, and Output Fluctuations." World Bank Institute, Washington, D.C. Processed.

Ahlumawia, Pavan. 2000. "Discriminating Contagion: An Alternative Explanation of Contagious Currency Crises in Emerging Markets." IMF Working Paper WP/00/14. International Monetary Fund, Washington, D.C.

Alba, Pedro, Amar Bhattacharya, Stijn Claessens, Swati R. Ghosh, and Leonardo Hernandez. 1999. "Volatility and Contagion in a Financially Integrated World: Lessons from East Asia's Recent Experience." In Gordon de Brouwer with Wisarn Pupphavesa, eds., *Asia Pacific Financial Deregulation.* London and New York: Routledge.

Arias, Eduardo Fernandez, Ricardo Hausmann, and Roberto Rigobon. 1998. "Contagion on Bond Markets: Preliminary Notes." Massachusetts Institute of Technology, Sloan School of Management, Cambridge, Mass. Processed.

Baig, T., and I. Goldfajn. 1998. "Financial Market Contagion in the Asian Crisis." IMF Working Paper WP/98/155. International Monetary Fund, Washington, D.C.

Banerjee, Abhijit. 1992. "A Simple Model of Herd Behaviour." *Quarterly Journal of Economics* 107:797–817.

Berg, Andrew, and Catherine Pattillo. 1999. "Are Currency Crises Predictable? A Test." *IMF Staff Papers* 46 (June):107–38.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. "A Theory of Fads, Fashion, Custom, and Cultural Changes as Information Cascades." *Journal of Political Economy* 1:992–1020.

Bordo, Michael D., and Antu P. Murshid. 2000. "The International Transmission of Financial Crises before World War II: Was There Contagion?" The Contagion Conference.

Bushee, Brian, and Christopher Noe. 1999. "Unintended Consequences of Attracting Institutional Investors with Improved Disclosure." Working Paper 00-33. Harvard Business School, Cambridge, Mass. Processed.

Calvo, Guillermo. 1998. "Understanding the Russian Virus (with Special Reference to Latin America)." University of Maryland, Department of Economics, College Park, Md. Processed.

Calvo, Guillermo, and Enrique Mendoza. Forthcoming. "Rational Herd Behavior and the Globalization of the Securities Market." *Journal of International Economics.*

Calvo, Sara, and Carmen Reinhart. 1996. "Capital Flows to Latin America: Is There Evidence of Contagion Effect?" In Guillermo Calvo, Morris Goldstein, and Eduard Hochreiter, eds., *Private Capital Flows to Emerging Markets after the Mexican Crisis.* Washington, D.C.: Institute for International Economics.

Caramazza, Francesco, Luca Ricci, and Ranil Salgado. 1999. "Trade and Financial Contagion in Currency Crises." IMF Working Paper WP/00/55. International Monetary Fund, Washington, D.C. Processed.

Chang, Roberto, and Giovanni Majnoni. 2000. "International Contagion: Implications for Policy." The Contagion Conference.

Choe, Hyuk, Bong-Chan Kho, and René Stulz. 1999. "Do Foreign Investors Destabilize Stock Markets? The Korean Experiences in 1997." *Journal of Financial Economics* 54(2):227–64.

Chuhan, Punam, Stijn Claessens, and Nlandu Mamingi. 1998. "Equity and Bond Flows to Asia and Latin America: The Role of Global and Country Factors." *Journal of Development Economics* 55:439–63.

Connolly, Robert, and Albert Wang. 1998. "International Equity Market Comovements: Economic Fundamentals or Contagion?" University of North Carolina, Kenan-Flagler Business School, Chapel Hill, N.C. Processed.

Corsetti, Giancarlo, Paolo Pesenti, and Nouriel Roubini. 1998. "What Caused the Asian Currency and Financial Crises? Part I: A Macroeconomic Overview." NBER Working Paper 6833. National Bureau of Economic Research, Cambridge, Mass.

Corsetti, Giancarlo, Paolo Pesenti, Nouriel Roubini, and Cédric Till. 1999. "Competitive Devaluations: A Welfare-Based Approach." NBER Working Paper 6889. National Bureau of Economic Research, Cambridge, Mass.

De Gregorio, José, and Rodrigo O. Valdés. 2000. "Crisis Transmission: Evidence from the Debt, Tequila, and Asian Flu Crises." The Contagion Conference.

Diamond, Douglas, and Philip Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91:401–19.

Diwan, Ishac, and Bernard Hoekman. 1999. "Competition, Complementarity, and Contagion in Asia." In Pierre-Richard Agenor, Marcus Miller, David Vines, and Axel Weber, eds., *The Asian Financial Crisis: Cause, Contagion, and Consequences.* Cambridge, U.K.: Cambridge University Press.

Dornbusch, Rudiger. 1999. "Brazil's Incomplete Stabilization and Reform." Massachusetts Institute of Technology, Department of Economics, Cambridge, Mass. Processed.

Dornbusch, Rudiger, Ilan Goldfajn, and Rodrigo Valdés. 1995. "Currency Crises and Collapses." *Brookings Papers on Economic Activity* 2:219–95.

Doukas, John. 1989. "Contagion Effects on Sovereign Interest Rate Spread." *Economic Letters* 29:237–41.

Drazen, Allan. 1999. "Political Contagion in Currency Crises." NBER Working Paper 7211. National Bureau of Economic Research, Cambridge, Mass.

Edwards, Sebastian. 1998. "Interest Rate Volatility, Capital Controls, and Contagion." NBER Working Paper 6756. National Bureau of Economic Research, Cambridge, Mass.

Eichengreen, Barry, Andrew Rose, and Charles Wyplosz. 1996. "Contagious Currency Crises." NBER Working Paper 5681. National Bureau of Economic Research, Cambridge, Mass.

Favero, Carlo A., and Francesco Giavazzi. 2000. "Looking for Contagion: Evidence from the 1992 ERM Crisis." The Contagion Conference.

Fischer, Stanley. 1998. "Reforming the International Monetary System." Draft paper prepared for the David Finch Lecture, Melbourne, Australia, presented November 9. Available at <http://www.imf.org/external/np/speeches/1998/110998.htm>.

Forbes, Kristin, and Roberto Rigobon. 1999. "No Contagion, Only Interdependence: Measuring Stock Market Comovements." NBER Working Paper 7267. National Bureau of Economic Research, Cambridge, Mass.

———. 2000. "Measuring Contagion: Conceptual and Empirical Issues." The Contagion Conference.

Frankel, Jeffrey, and Sergio Schmukler. 1998. "Crisis, Contagion, and Country Fund: Effects on East Asia and Latin America." In Reuven Glick, ed., *Managing Capital Flows and Exchange Rates: Perspectives from the Pacific Basin.* New York: Cambridge University Press.

Froot, Kenneth A., Paul G. J. O'Connell, and Mark S. Seasholes. 2000. "The Portfolio Flows of International Investors, I." The Contagion Conference.

Furman, Jason, and Joseph Stiglitz. 1998. "Economic Crises: Evidence and Insights from East Asia." *Brookings Papers on Economic Activity* 2:1–135.

Garber, Peter. 1998. "Derivatives in Capital Flows." NBER Working Paper 6623. National Bureau of Economic Research, Cambridge, Mass.

Gelos, R. Gaston, and Ratna Sahay. 2000. "Financial Market Spillovers in Transition Economies." The Contagion Conference.

Gerlach, Stefan, and Frank Smets. 1995. "Contagious Speculative Attack." *European Journal of Political Economy* 11:5–63.

Glick, Reuven, and Andrew Rose. 1998. "Contagion and Trade: Why Are Currency Crises Regional?" NBER Working Paper 6806. National Bureau of Economic Research, Cambridge, Mass.

Goldberg, Linda, Gerard Dages, and Daniel Kinney. 2000. "Lending in Emerging Markets: Foreign and Domestic Banks Compared." The Contagion Conference.

Goldfajn, Ilan, and Rodrigo O. Valdés. 1997. "Capital Flows and the Twin Crises: The Role of Liquidity." IMF Working Paper WP/97/87. International Monetary Fund, Washington, D.C.

Goldstein, Morris, Graciela Kaminsky, and Carmen Reinhart. 2000. *Assessing Financial Vulnerability: An Early Warning System for Emerging Markets.* Washington, D.C.: Institute for International Economics.

Hawkins, John, and Philip Turner. 2000. "International Financial Reform: Some Issues." The Contagion Conference.

Huh, Chan, and Kenneth Kasa. 1997. "A Dynamic Model of Export Competition, Policy Coordination, and Simultaneous Currency Collapse." Pacific Basin Working Paper Series PB97-08. Federal Reserve Bank of San Francisco, San Francisco, Calif.

Jeanne, Olivier. 1997. "Are Currency Crises Self-Fulfilling?" *Journal of International Economics* 43:263–86.

Kaminsky, Graciela, and Carmen Reinhart. 1998a. "Financial Crisis in Asia and Latin America: Then and Now." *American Economic Review* 88(May, *Papers and Proceedings*):444–48.

———. 1998b. "On Crises, Contagion, and Confusion." George Washington University, Economics Department, Washington, D.C. Processed.

Kaminsky, Graciela, and Sergio Schmukler. 1999. "On Booms and Crashes: Stock Market Cycles and Financial Liberalization." George Washington University, Economics Department, Washington, D.C. Processed.

Kaminsky, Graciela, Richard Lyons, and Sergio Schmukler. Forthcoming. "Economic Fragility, Liquidity, and Risk: The Behavior of Mutual Funds during Crises." *Journal of International Economics.*

Kim, Jungshik, and Shang-Jin Wei. 1997. "The Big Players in the Foreign Exchange Market: Do They Trade on Information or Noise?" NBER Working Paper 6256. National Bureau of Economic Research, Cambridge, Mass.

Kim, Woochan, and Shang-Jin Wei. 1999. "Foreign Portfolio Investors before and during a Crisis." NBER Working Paper 6968. National Bureau of Economic Research, Cambridge, Mass.

Kindleberger, Charles. 1989. *Manias, Panics, and Crashes.* Rev. ed. New York: Basic Books.

Kodres, Laura, and Matthew Pritsker. 1998. "A Rational Expectations Model of Financial Contagion." Finance and Economics Discussion Paper 9848. Board of Governors of the Federal Reserve System, Washington, D.C.

Masson, Paul. 1998. "Contagion: Monsoonal Effects, Spillovers, and Jumps between Multiple Equilibria." IMF Working Paper WP/98/142. International Monetary Fund, Washington, D.C.

Park, Yung Chul. 1998. "Financial Crisis in Korea and Its Lessons for Reform of the International Financial System." In Jan Joost Teunissen, ed., *Regulatory and Supervisory Challenges in a New Era of Global Finance.* The Hague: FONDAD.

Park, Yung Chul, and Chi-Young Song. 1999. "East Asian Financial Crisis: One Year After." *IDS Bulletin* 30(1):93–103. Institute of Development Studies, University of Sussex, Brighton, U.K.

———. 2000. "Financial Contagion in the East Asian Crisis: With Special Reference to the Republic of Korea." The Contagion Conference.

Perry, Guillermo, and Daniel Lederman. 1998. "Financial Vulnerability, Spillover Effects, and Contagion." World Bank, Chief Economist's Office, Latin America, Washington, D.C. Processed.

Pritsker, Matt. 2000. "The Channels for Financial Contagion." The Contagion Conference.

Radelet, Steven, and Jeffrey Sachs. 1998a. "The East Asian Financial Crisis: Diagnosis, Remedies, Prospects." *Brookings Papers on Economic Activity* 1:1–90.

———. 1998b. "The Onset of the East Asian Currency Crisis." NBER Working Paper 6680. National Bureau of Economic Research, Cambridge, Mass.

Sachs, Jeffrey, Aaron Tornell, and Andres Velasco. 1996. "Financial Crises in Emerging Markets: The Lessons from 1995." *Brookings Papers on Economic Activity* 1:147–215.

Scharfstein, David, and Jeremy Stein. 1990. "Herd Behavior and Investment." *American Economic Review* 80:465–79.

Schinasi, Garry J., and R. Todd Smith. 2000. "Portfolio Diversification, Leverage, and Financial Contagion." The Contagion Conference.

Shiller, Robert. 1995. "Conversation, Information, and Herd Behavior." *American Economic Review* 85 (May, *Papers and Proceedings*):181–85.

Stiglitz, Joseph E., and Amar Bhattacharya. Forthcoming. "Underpinnings for a Stable and Equitable Global Financial System: From Old Debates to a New Paradigm." *Annual World Bank Conference on Development Economics 1999.* Washington, D.C.: World Bank.

Stulz, René M. 1999. "International Portfolio Flows and Security Returns." Working Paper 99:3. Ohio State University, Charles A. Dice Center for Research in Financial Economics, Columbus, Ohio. Available at <http://www.cob.ohio-state.edu/fin/faculty/stulz>.

Tan, Jose A. 1998. "Contagion Effects during the Asian Financial Crisis: Some Evidence from Stock Price Data." Pacific Basin Working Paper Series PB98-06. Federal Reserve Bank of San Francisco, San Francisco, Calif.

Valdés, Rodrigo. 1997. "Emerging Markets Contagion: Evidence and Theory." Massachusetts Institute of Technology, Economics Department, Cambridge, Mass. Processed.

Van Rijckeghem, Caroline, and Beatrice Weder. 2000. "Financial Contagion: Spillovers through Banking Centers." The Contagion Conference.

Wermers, Russ. 1995. "Mutual Fund Herding and the Impact on Stock Prices." *Journal of Finance* 54(April):581–622.

Wolf, Holger. 1999. "International Asset Price and Capital Flow Comovements during Crisis: The Role of Contagion, Demonstration Effects, and Fundamentals." Paper presented at the World Bank/IMF/WTO conference on "Capital Flows, Financial Crises, and Policies," April 15–16, Washington, D.C.

Weak Links in the Chain: A Diagnosis of Health Policy in Poor Countries

Deon Filmer • Jeffrey S. Hammer • Lant H. Pritchett

Recent empirical and theoretical literature sheds light on the disappointing experience with implementation of primary health care programs in developing countries. This article focuses on the evidence showing two weak links in the chain between government spending for services to improve health and actual improvements in health status. First, institutional capacity is a vital ingredient in providing effective services. When this capacity is inadequate, health spending, even on the right services, may lead to little actual provision of services. Second, the net effect of government health services depends on the severity of market failures—the more severe the market failures, the greater the potential for government services to have an impact. Evidence suggests that market failures are the least severe for relatively inexpensive curative services, which often absorb the bulk of primary health care budgets. A companion paper, available from the authors (see p. 219), offers a perspective on how government funds can best be used to improve health and well-being in developing countries. It gives an alternative view of appropriate public health policy, one that focuses on mitigating the characteristic market failures of the sector and tailoring public health activities to the government's ability to deliver various services.

The concept called primary health care, or health for all, enshrined at an international conference at Alma Ata (now Almaty, Kazakhstan) in 1978, has dominated much of the discussion of health policy in developing countries for the past two decades. The broad health problems this concept encompasses are pressing, and the solutions it calls for seem obvious. Yet in many developing countries the public budget for health is principally absorbed by public hospitals staffed by doctors expensively trained at public expense who use costly medical technologies to treat conditions of the urban elite, while in those same countries children die from diseases that could have been treated for a few cents or avoided altogether with basic hygienic practices. In developing countries in 1995, more than 9 million children under five

The World Bank Research Observer, vol. 15, no. 2 (August 2000), pp. 199–224.

years old—more than the entire population of Sweden or Zambia—died deaths that could have been avoided.[1]

Individual examples, however, show that primary health care efforts can be successful even in very poor regions. Kerala, a state in India with annual per capita income of only $1,254 in purchasing power parity dollars, has an infant mortality rate of only 31 per 1,000 live births (Agrawal and others 1996; International Institute for Population Studies 1995). This rate is not only 40 percent lower than that in Punjab, another Indian state with twice Kerala's income, but 35 percent lower than that in Brazil, which has more than four times Kerala's per capita income (Macro International 1997; World Bank 1999). Infant mortality in Shanghai, China, is lower than in Manhattan, and the recorded infant mortality rate of 16 in Jamaica is lower than that of African-Americans in the United States. Ceara, one of the poorest states in Brazil, reduced infant mortality by 36 percent in just a few years through an aggressive government program (Tendler and Freedheim 1994).

This combination of experiences has led to a strong consensus among public health specialists who focus on developing countries. They argue that the existing allocation of health expenditures toward curative care in secondary and tertiary facilities, such as hospitals and clinics to which patients are referred, is inappropriate and that a reorientation of government efforts toward primary health care would bring both health gains and cost savings. In this consensus primary health care is typically defined by what it is *not:* it is neither secondary nor tertiary curative care but could be *all* other activities related to health, from nutrition, to sanitation, information, and education, to clinic-based curative care. Even more ambitious definitions include empowerment and social equality (Decosas 1990).

Although the images and statistics that motivate primary health care appear compelling, the gains have rarely been demonstrated in practice. Individual initiatives, such as the expansion of immunization campaigns or the recent efforts to combat river blindness, have been effective in a single region or for a single disease, but the data do not typically show that other aspects of public health care have much impact on health. The implementation of primary health care programs has created a new set of images: empty rural health clinics without staff, drugs, or working equipment; poor people bypassing free primary public clinics to pay for services from private providers; government-supplied drugs for sale on the black market. This disappointing experience raises the question: what was missing from the seemingly compelling logic? There appear to be several weak links in the chain that prevent public primary health care from being delivered to more people. We believe that economists bring to bear two perspectives on this question that are useful for understanding health policy: choices and incentives.

First, the expected impact of primary health care was too often calculated as if health status were entirely a technical affair and individuals were the passive recipients of government action. But individuals are guided by their own knowledge and

resources in judging the quality of their health care (and that of their children). Incorporating their choices into the analysis can change both the expected overall impact of these programs and the importance of various actions, because service delivery and overall effectiveness both depend on the demand for specific services, the price of the services, and the existing (and potential) supply in the private sector.

Second, primary health care advocates assumed that the public sector could deliver whatever the government (or some international forum) decided *ought* to be delivered. In practice, the quality of public health services has ranged from excellent to truly horrific. While an ideal, well-run network of community workers and rural health clinics might have a dramatic effect on health status, the real policy issue is what the public sector is actually capable of providing. Often, health service failures result from a systemic mismatch between the traditional civil service incentive structure and the tasks required in the health sector.

A Simple Framework for Analyzing Public Spending on Health

In general, public spending influences health status by lowering the effective price of health-enhancing inputs (whether information on food cleanliness or a heart transplant). How it does so depends on four distinct mechanisms:

- *Composition of public spending.* The effect of an increase in total public spending on health depends on how that increase is allocated across health inputs. An equal increase in spending on all inputs will have a very different effect from an increase in spending on only the most effective interventions.
- *Output of the public sector.* The next issue is the magnitude of effective health services created by the public spending. Whether a government decides to build a clinic, spray for malaria, mobilize community outreach workers, or buy X-ray machines, it can be more or less effective at translating that expenditure into a real supply of health services. This efficacy will have a country-specific component common to all activities, as well as an activity-specific component, since government might be better adapted to performing some activities than others.
- *Net impact of public sector supply on overall consumption.* Even when government funds are available to provide the service in question, one must assess whether the change in the effective price faced by consumers will translate into increased overall consumption. Public spending on a particular service may not be cost-effective in improving health if additional consumption in the public sector crowds out—in whole or in part—equally effective services from nongovernmental providers. The size of this crowding-out effect depends on how individual demand and private sector supply respond to changes in the price, travel time, convenience, or quality of services available.

- *The health production function.* Different health inputs are more or less effective in improving health in ways determined by biological and medical facts. This health production function—that is, determinations about which treatments are effective in eliminating which cancers, which vaccines are potent over what period, how micronutrients affect susceptibility to diseases—is the domain of health care professionals. Economists typically prefer to remain agnostic about the particulars of the production function, and in many cases sensible recommendations about public policy need not inquire into the production function.

Within this framework, the consensus argument is that increases in public spending on primary health care are effective in improving aggregate measures of health status, while curative services at secondary and tertiary levels are not.[2] This view can rationalize an increase in funds for primary health care as well as a reallocation of the health budget toward primary health care activities. The argument, however, typically jumps directly from the assumption that if the government spends on the right things, patients will receive the right things. Primary health care advocates rely on evidence showing that when evaluated as interventions, primary health care elements of the health production function are typically more cost-effective than secondary or tertiary activities. But the actual impact of public spending is the product of all four terms described above: allocation of the budget, efficacy of the public sector, market impact on consumer demand for services, and the actual effect of the health services on health. If any one of these is low, the total impact will be low, and the middle two are the weak links.[3]

Has Primary Health Care Been Effective at the National and Local Levels?

Has public spending on health and, more particularly, on primary health care promoted good health? If so, there should be empirical regularities at both the national and local levels. First, given a level of total health expenditures, more spending on primary health care activities and greater access to primary health care services should be associated with lower aggregate mortality. Second, at the local level (household, village) greater access to primary health care facilities should reduce mortality. Third, projects that develop primary care facilities should reduce mortality. None of these regularities finds much support in the data.

It might seem odd that we do not review the literature evaluating primary health care per se. As one recent review of that literature highlights, however, the literature does not permit such an evaluation. Fox (1995:2) analyzed the 87 articles published since 1980 in four major health journals (*Health Policy, Health Policy and Planning,*

International Journal of Health Services, and *Social Science and Medicine*) that included the words "primary health care" in the Medline subject category:

> While there are questions about parts of PHC [primary health care] (e.g. cultural acceptability of community participation ideals, selection processes for PHC workers, community financing critiques, the capture of program planning or implementation by funding agencies, or the ability to have health for all in socially unjust societies), there was no serious questioning in the literature of PHC as a desirable way for Ministries of Health to spend their money. . . . Most of the program evaluations or topic evaluations don't show great technical expertise in evaluation methodology being applied to PHC, but a fair number of articles discuss how such an evaluation would be done, or why it would be difficult.

A similar problem affects the empirical evaluation of "health status." The World Health Organization (WHO) defines health as "a state of complete physical, mental, and social well-being, and not merely the absence of disease or infirmity" (WHO 1988). Although this is an attractive definition, it is subjective and hard to assess. Mortality is easy to measure, whereas morbidity is not, but (fortunately for people, but unfortunately for research) mortality is rare (except among the very young and very old). Moreover, because death occurs rarely, and only once for each person, it is difficult to study at the individual level. So even though interest focuses on health over an individual's lifetime, empirical studies tend to view infant (or child) mortality or life expectancy as proxies for health status and to rely on mortality or life expectancy rates at an aggregate level (district, province, and country) for comparisons.[4]

Comparisons at the National Level

Cross-national studies of health status have come to a fair consensus on two points. First, socioeconomic characteristics explain nearly all of the variation in mortality rates across countries. A recent econometric study of mortality among children under age five shows that virtually all the cross-national variation in child mortality can be explained by six variables: average gross domestic product (GDP) per capita, a measure of the distribution of income, the level of female education, a dummy variable for predominantly Muslim countries, an index of ethnolinguistic diversity, and a set of five dummy variables for regions (Filmer and Pritchett 1999).

Second, total public spending on health has had much less impact on average health status than one might have expected and certainly less than one could have hoped for. Although the lack of data on public spending has, until recently, limited direct examination of the issue, Musgrove (1996:44) summarizes studies of the effect of public spending on health as follows: "Multivariate estimates of the determi-

nants of child mortality give much the same answer [as his results on life expectancy]: income is always significant, but the health share in GDP, the public share in health spending, and the share of public spending on health in GDP never are." A recent paper on global health from the Population Reference Bureau downplays the role of aggregate spending, stating that "these resources, however, usually do not adequately measure the *effectiveness* or availability of services" (Ratzan, Filerman, and LeSar 2000:36, emphasis added).

Other recent studies support these findings. Using instrumental variables to account for data and endogeneity problems, Filmer and Pritchett (1999) find that public expenditure on health as a share of GDP is a small and statistically insignificant determinant of child mortality. Estimates find that doubling public spending from 3 to 6 percent of GDP would improve mortality by only 9–13 percent. Bidani and Ravallion (1997) show that public spending has a large impact on the health status of the poor, but they estimate the effect of public spending on aggregate health status (of the poor and nonpoor taken together) to be quite small.

We say the consensus is "fair" because there are some exceptions to these two points. First, Anand and Ravallion (1993) could not reject the restriction that average income affected health status only insofar as it affected the level of poverty, highlighting the fact that increased income among the poor is the most effective variable in improving health. The very high correlation between average income and poverty still implies, however, that a country's average income explains most of the variation in health status. Second, an influential study (Preston 1980), based on data from 1940 to 1970, emphasized the low explanatory power of socioeconomic variables. Studies using more recent data, however, including one by Preston (1986), who relied on data from the 1970s, are unanimous about the high explanatory power of socioeconomic basics such as average income and female education.

The cross-national evidence has always been absent or ambivalent on whether health status is improved by a greater commitment to or greater spending on primary health care (or both). First, despite the power of socioeconomic conditions to determine health status, there are still outliers, such as Kerala in India, Sri Lanka, and Costa Rica, whose achievements are potentially replicable. Although these examples suggest possibilities, it was never very clear whether the success stemmed directly from characteristics of their health systems or was instead a social or political phenomenon. Kunstadter (1985:234), participating at a seminal conference that affirmed support for primary health care (at least of the selective variety) on the basis of case studies of the outliers, commented:

> The four case studies [China, Costa Rica, Kerala State, and Sri Lanka] involve societies in which low mortality has been reached without high per capita income. Situations in which low income continues to be associated with high mortality or high income is associated with high mortality were not

considered, nor have we searched systematically for other societies in which relevant social characteristics of the four successful cases are repeated, to see what happened to mortality. Thus the policy prescriptions are relatively weak.

Moreover, although China's barefoot doctors are famous, it is not obvious that their success can be attributed to a primary health care strategy—nor even that it was related to medical care. The largest declines in infant and child mortality occurred in the 1950s and the first half of the 1960s, before the introduction of the barefoot doctors in late 1965. Further, when the program was abandoned (ironically, at about the time of the Alma Ata conference), health status did not deteriorate (Liu and others 1995). In 1986 Sri Lanka spent 70 percent of its public monies on hospitals, substantially higher than the 56 percent average for comparable countries in South Asia (Griffin 1992). Our regressions for this article suggest that when either the fraction of the budget spent on "local health services" or the "fraction of population with access to local health services" is added to a regression of mortality on income and other country socioeconomic characteristics, it has no statistically or empirically significant impact (Filmer, Pritchett, and Hammer 1998).[5] Although this is weak evidence *against* the effectiveness of primary health care, we know of no published cross-national evidence that lends *support* to primary health care.

Outcomes at the Local Level

A second empirical regularity that would support primary health care would be evidence that the availability of primary-level health facilities or community health workers had a demonstrable impact on the local health status of individuals and communities. Unfortunately, the empirical results about the effect on health status of proximity to hospitals, doctors, and, in particular, public sector clinics, health centers, and rural health workers are, at best, mixed.

A major problem in assessing this evidence is that governments may have systematically located clinics in underserved areas where health status is worst. As a result, the comparison of health status in localities with and without clinics would understate the true impact of such clinics. Conversely, if the government places clinics in villages where the demand is greatest, these sites may be located in areas where health would have been good in any case, and comparisons of sites with and without a facility would overstate the impact of clinics.

Recent studies assess the effect of access to services on child or infant mortality, using methods that take into account the bias resulting from the placement of facilities. Frankenberg (1995) controls for placement effects by comparing randomly matched births from two different age cohorts in a village in Indonesia. She finds that the presence of a maternity clinic or a doctor reduced mortality but that the

presence of a health worker other than a doctor *increased* the probability of death (although this finding was statistically significant only at the 10 percent level). Pitt, Rosenzweig, and Gibbons (1993) compare matched districts in Indonesia and find that villages in a district served by a health center had higher mortality rates, whereas those with a family planning clinic had lower mortality rates; in both cases, the findings are statistically insignificant. When she does not control for selective placement, Frankenberg (1995) finds that the presence of more maternity clinics and health workers (insignificantly) reduced mortality, while more doctors (insignificantly) increased mortality. In uncorrected estimates, Pitt, Rosenzweig, and Gibbons (1993) find that the presence of both health centers and family planning clinics raised mortality rates (insignificantly in the case of health centers).

There are many studies that do not control for selective placement. After controlling for the potential endogeneity of facility usage, Panis and Lillard (1994) find that delivering a baby within a health care institution in Malaysia (the likelihood of which increases with the availability of such facilities) reduced the probability that the baby would subsequently die; puzzlingly, they also find that prenatal care (insignificantly) increased the probability that the child would subsequently die. Benefo and Schultz (1996) find that proximity to a clinic reduced child mortality in Côte d'Ivoire but appeared to increase mortality (statistically insignificantly) in Ghana. Lavy and others (1996) find that proximity to health facilities significantly increased mortality in rural Ghana (although the availability of child services in the closest clinic was significantly positively related to survival). In Malaysia DaVanzo (1984) finds that distance to medical care was (insignificantly) related to infant mortality (conditional on birthweight) but that low birthweight was correlated with the distance to care. Hammer, Nabi, and Cercone (1995) find that public medical facilities in Malaysia were unrelated to mortality, whereas Hossain (1989) reports that the presence of a dispensary and a family planning clinic lowered mortality in Bangladesh.

Rosenzweig and Schultz (1982) find that rural health posts, municipal public and private clinics, dispensaries, and mobile care units were not significantly related to child mortality in rural Colombia. In urban areas, hospitals, clinics, and family planning centers tended to reduce mortality, but this result was not consistent across groups sorted by the age of the mother. Rosenzweig and Wolpin (1982) use data from rural India to find that villages in a district with a family planning clinic and those with a dispensary were associated with lower child mortality but that villages with any "other health facility (health centers, nursing homes, etc. . . .)" were associated with higher mortality. The large National Family Health Survey of India shows no relation between health centers or subcenters and child mortality (World Bank 1998a).

Sastry (1995) finds that the number of general health facilities is associated with higher mortality in northeastern Brazil but with lower mortality in southern and southeastern Brazil (both effects are statistically insignificant). In a study of census

districts in Brazil, Paes de Barros, Da Costa, and Mendoça (1998) find that the availability of public health personnel had no effect on mortality.

Overall, the econometric evidence that clinics have *any* effect on health is mixed. Even where such evidence exists, the impact is too small to explain much of the variation in health status.

Evaluating the Effect of Health Care Programs on Child Mortality

In 1977 a very intensive maternal and child health and family planning program was introduced in a set of treatment villages in the Matlab region of Bangladesh (with a nearby set of villages serving as comparators). In the treatment area mothers and children were visited every 15 days by a female health worker who provided guidance on family planning. Detailed records were kept in both areas. Although mortality among children fell (Muhuri and Preston 1991), analysts exploiting the controlled design of this project attribute the decline almost exclusively to measles immunization (Koenig, Fauveau, and Wojtyniak 1991; Menken and Phillips 1990).

A separate project that assessed the impact of health care services for children in Narangwal, India, found that infant mortality dropped 6 percentage points over the three years of the project (from 96 per 1,000) in the treatment area compared with a 1 percentage point increase in the control area (from 107 per 1,000), but the difference was insignificantly different from zero (Taylor and Singh n.d.). The change in the project villages resulted from a 10 percentage point increase in perinatal mortality, a 4 percentage point increase in neonatal mortality, and a 16 percentage point decrease in postneonatal mortality. In the control villages, perinatal mortality rose only 1 percentage point, neonatal mortality fell 21 percentage points, and postneonatal mortality increased 49 percentage points. The last figure leads the authors to question the resulting 1 percentage point increase reported for overall infant mortality. The authors' explanation is that this huge fluctuation was "due to incorrect age classification, . . . since in the control villages . . . investigators who had no access to exact birth information needed to rely on the age at death given by the mother of the family" (p. iv.D.7).

These empirical results suggest that enhancing health outcomes is not simply a matter of providing additional funds or increasing access to primary health care services and facilities. The framework discussed earlier provides two likely explanations for the negligible impact of public spending. Each of these explanations leads to different policy implications. First, the impact of primary health care provision depends on the effectiveness of the service provided. If this is the problem, the solution is an improvement in the quality of the service. The question then becomes: what policy levers are available to increase quality? Second, the impact of the service depends on individual choice and the market for health, that is, private demand may vary by disease condition and the response by private suppliers to public interven-

tion. The question in this case is, what are the factors that affect private supply and individual demand?

Public Sector Spending and the Creation of Effective Health Services

Primary health care may have little impact on health status not because such care is unimportant but because, in practice, the efficacy of government health interventions may be low. Readers in industrial countries without personal experience may find it difficult to appreciate just how poor the quality of public sector services can be. Anecdotal reports describe situations marked by corruption and mismanagement. In one low-income country, for example, a prominent newspaper accused the health ministry of misappropriating $50 million of donor financing. The next day the ministry accused the newspaper of exaggeration for failing to make it clear that this $50 million was misappropriated over a period of three years, not in a single year as the newspaper report implied.

In a second case, a client survey of women who had delivered a child in the past two years at rural health centers in the Mutasa district of Zimbabwe (Mtemeli 1994) listed the most frequently cited disadvantages of giving birth in an institution. These were "ridiculed by nurses" for not having baby clothes (22 percent), "maternity fees" (16 percent), "nurses ordered mothers to wash used linen soon after delivery" (16 percent), and "nurses hit mothers during delivery" (!) (13 percent). Interestingly, when the nurses themselves were asked why they thought mothers did not deliver in health institutions, they cited, first, the distance to the facility and transport problems (20 percent); second, inability to pay (14 percent); and, third, *harassment by nursing staff or fear of nurses* (11 percent).

Moreover, in nearly every country one can find rural health clinics that have no drugs, although government- (or donor-) financed medicines are readily available on the black market. For example, more than 70 percent of the government supply of drugs disappeared in Guinea in 1984 (Foster 1990). Various studies in Cameroon, Tanzania, and Uganda estimated that about 30 percent of publicly supplied drugs was misappropriated; in one case as much as 30 to 40 percent of the public supply was "withdrawn for private use" by staff (World Bank 1994a).

Evidence on Quality

Alderman and Lavy (1996) recently reviewed several empirical studies that examine the link between the quality of and the demand for public facilities; among these studies were Akin, Guilkey, and Denton (1995); Lavy and Germain (1994); Lavy and others (1996); Mwabu, Ainsworth, and Nyamete (1993); and Thomas, Lavy,

and Strauss (1996). Even though the measures of quality are not always satisfactory, the findings confirm that demand is responsive to service characteristics. An example of the problematic nature of some of the measures is that the absence of various types of drugs is often used to indicate inadequacies in public health delivery. Shortages, however, could be caused by high levels of demand, and inferences about the causal relationship are difficult to draw (a problem generally acknowledged by the authors cited above). In addition, there may be important discrepancies between de facto and de jure measures of quality, a result highlighted by Thomas, Lavy, and Strauss (1996), who show the different effects on health outcomes of the actual number of staff members compared with the official number.

Both the ineffectiveness of low-level public sector health clinics and the element of individual choice are highlighted in the phenomenon called "bypassing," which occurs when people choose to bypass the closest public facility in favor of either more costly private facilities or higher-level public facilities. Because detailed information on the health choice behavior of individuals as well as on all potential sources of supply is needed to understand bypassing, there are few empirical studies of this phenomenon. Analyses document large amounts of bypassing in Sri Lanka, where detailed surveys of health care supply and demand were collected (Akin and Hutchinson 1999; Samrasinghe and Akin 1994). Samrasinghe and Akin (1994) show that 31 percent of all episodes of illness were self-treated without medical consultation. Of those people who did seek treatment, only 42 percent were treated at the closest facility—either a predominantly private ayurvedic, or traditional, practice or a government-run, low-level health facility offering a Western approach to care; 58 percent did not go to the nearest facility. Of the non-ayurvedic sources, the public facilities (the type that primary health care would promote) are most frequently bypassed. Most individuals who bypass either a minor or a major public facility do so to visit a private Western facility.

The authors find that the bypassed public facilities had fewer doctors, nurses, and services and less equipment than public facilities that were not bypassed. In contrast, the private Western facilities that were bypassed had more doctors, nurses, and services and higher levels of care. Although this result might seem paradoxical, it is consistent with sophisticated health-seeking behavior. Because prices tend to be lower in the public facilities, patients will bypass high-quality, but expensive, private facilities in favor of public ones when their condition is not serious or quality is not important. For serious conditions, however, or when quality of service is important, individuals are willing to pay—in terms of both time and fees—for higher-quality care.

Bypassing can lead to low utilization of available public facilities. A survey (PIEDR 1994:vi) of a rural area of Punjab Province, Pakistan, found that although the physical infrastructure of rural primary health care was in place, "[o]nly about 5 percent of the sick children were taken for treatment to primary health care facilities; half were taken to private dispensers, and another quarter to private . . . doctors. Around 95

percent of deliveries took place at home." Roughly the same percentage of respondents sought treatment from a public rural health facility (5.2 percent) as from a "quack" (4.9 percent) (PIEDR 1994:35). The decision to bypass the public facility did not reflect long queues at those facilities. On the contrary, the typical rural health center saw only about 30 patients a day and the typical basic health center only 11 patients a day—far below their usual capacity, as rural health centers employed an average of 8 workers and basic health units, 5. Two atypically busy rural health centers attracted and serviced an average of more than 450 patients a day.

A study of health centers in Indonesia also found low usage of public facilities. Annual caseloads were low even for facilities located near large local populations (World Bank 1994b). On the basis of detailed case studies, this study identified two principal reasons for the low usage. First, many public facilities were short of equipment, drugs, and appropriate health workers. Second, and more important, detailed assessments on the way public health clinics operated showed that poorly functioning facilities contributed to patients' decisions to seek medical care elsewhere. Respondents in one case study made it clear that public facilities were of low quality. "They were confident that they could get considerate and unrushed care in a pleasant and informal setting in the private practice of doctors, bidans [midwives] and nurses" (World Bank 1994b:10).

Quantitative and focus group techniques in El Salvador found remarkably similar results (Lewis, Eskeland, and Traa-Valerezo 1999:26–27). Respondents consistently complained about the low quality of public health units, especially compared with the services provided by health centers and hospitals. Typical complaints about the public health units were: "Health posts operate only twice a week. Consultation is only until noon. The doctor is not always there. Sometimes only the nurse assistant is present. Waiting time is three hours on average. Only those who arrive by 8 get a consultation." In one locality, Potrero Sula, a respondent is quoted as saying, "The [public health] post here is useless because there is no doctor or nurse, and it is only open two days a week until noon." In contrast, for example, a respondent in El Pinar described the private health center at La Palma as "a little hospital with very good services. It is well equipped. The fee is only 3 colones for consultation and sometimes medication."

The study in El Salvador assessed the effect of intervention at the very lowest level—that of the health worker who lives in the community and provides simple medical services but refers seriously ill patients to larger facilities. Focus group respondents had very little use for these public health workers, and regression analysis found that they had little or no demonstrable effect on individuals' decisions about where to seek medical treatment.

Even when medical care in the public sector is of reasonable quality, it may be tremendously inefficient. A stark example of the kinds of inefficiency possible in the public sector comes from a detailed study of costs and expenditures in a public hos-

pital in the Dominican Republic, measured through careful observations on use of time (Lewis, La Forgia, and Sulvetta 1996). The study showed that although spending on personnel constituted 84 percent of total recurrent spending, actual staff costs for treating patients were only 19 percent of total costs. Gross inefficiency was identified as the cause of this huge discrepancy, an inefficiency that could be explained by the incentives, or lack thereof, facing personnel. There was no accountability for physicians or nurses, no rewards for extraordinary performance, and no punishment for inadequate or nonexistent performance. Salaries were low and undifferentiated. There was no management control over staff and essentially no returns to effective management. Clearly, the private sector also has problems, but the profit motive acts as a powerful incentive for efficiency.

How Can the Public Sector Deliver High-Quality Services?

Even strong advocates of primary health care would agree that governments do not always provide high-quality health services cost-effectively. Conversely, even the most ardent critic of government acknowledges that there are admirable and well-functioning health facilities and agencies in the public sector. Is failure, then, the result of resource constraints, ignorance, and mistakes, or is it a systematic and expected result of (dis)incentives created by institutional and organizational arrangements? If one is convinced that existing public sector problems are easily remediable through larger budgets, earmarked inputs, additional training, or technical assistance of various kinds, there is no reason to back away from public provision as a strategy. But if one is convinced that the failures are endemic and intrinsic to the especially low capability of the public sector, then perhaps the entire strategy for delivering health services needs rethinking.

PAY, EMPLOYMENT, AND PERFORMANCE IN THE HEALTH SECTOR. There is no one right answer about capability in the public health sector. Some public agencies provide high-quality, cost-effective health care. Others are capable of spending unlimited amounts of resources with no health gains. Two tough policy questions must be addressed. First, are the conditions in place for effective public provision of particular health services? And second, if not, is it possible to establish those conditions within a reasonable time horizon? Although some critics have perhaps been overly pessimistic about government capacity, supporters of publicly provided clinical services have paid far too little attention to the first question and have been entirely too sanguine about the second.

The feasibility of linking pay and employment to work effort in the public sector is not a new issue. It is generally recognized that the more essential and the less easily observable is individual effort, the greater is the importance of linking pay and performance (Milgrom and Roberts 1992). Workers in situations where effort is easily

observed and monitored are usually paid wages or salaries, whereas those whose output is observed but whose effort is not, such as salesmen, are usually paid on the basis of outcomes. In addition to the level of pay, there is the question of continued employment. Where output is crucially linked to individual performance and there can be little tolerance for deviations from high quality, continued employment is generally linked to performance. Observation of pay and employment across the public and private sectors tends to reinforce this position. To illustrate, table 1 shows a matrix of jobs and the degree to which pay and employment tend to be related to performance.

Among private sector professionals (lawyers, say, or doctors in medical practices), both pay and employment tend to be highly related to performance; in traditional public sector organizations, neither is. Such an observation is perhaps a commonplace. Yet rarely have policymakers or economists questioned whether the public sector is the most appropriate provider of clinical services, given that outsiders can observe neither health efficacy nor client treatment.

One reason for this lack of inquiry is that sometimes the most obvious and seamlessly working features of a system are invisible and are taken for granted when the system is functioning well. As a consequence, when one attempts to extrapolate from one set of social, legal, and political conditions to another, very different, set, the key features may be missed. For example, why will local health workers do the right thing even though there is no disciplining device of consumer choice? They are underpaid, and there are no effective institutional controls or legal restraints (for example, malpractice suits). The common (implicit) answer is often something like "because they are trained health professionals." Indeed, in well-functioning systems this is most likely the right answer. Doctors and nurses do not perceive themselves as performing for the money, or because of threats or close scrutiny, but out of professional pride and affiliation. The underlying factors of compensation and punishment are invisible not because they are weak or absent but precisely because they are so strong and effective that gross deviations from appropriate behavior are rare, and hence the need

Table 1. *Various Types of Employment and Compensation Schemes*

Degree to which employment is linked to performance	*Degree to which pay is linked to performance*		
	Low	*Medium*	*High*
Low	Traditional civil service arrangements (postal workers, administrators)	Large, stable organizations	Piece rate (agricultural workers, salesmen, contract workers)
High	"Up or out" organizations (U.S. military)	Most private sector organizations	Professionals (law firms, medical practices)

Source: Authors' summary.

to invoke explicit punishment is similarly rare. When the underlying control mechanisms are weak or nonexistent, however, professionalism is not a powerful enough inducement.

The second incentive issue involves the proper mix of public and private providers. Such a mix is advantageous because mobility between the two sectors can act as a disciplining device on the behavior of individuals in the public sector. Doctors working in a public hospital when there is a large and effective private sector do not want to damage their reputations by performing noticeably worse than their colleagues. With either too little or too much interchange, however, this regulating mechanism will cease to be effective. The former effect may be due to a small private sector, the latter to a large private sector with professionals working in both sectors and with no standards or monitoring of the public part of their practice.

LEARNING FROM SUCCESS. What are the lessons from those instances in which the public sector is effective at improving health outcomes? Close examination points to the importance of social, political, and institutional factors in motivating effective performance from health workers. However, although the declaration at Alma Ata stated that one manifestation of these factors, "community participation," was a key feature of the primary health care strategy, it appears to be more the exception than the rule.

The role of social and political factors in generating the effective performance of government agencies in Kerala is well described in Heller (1996). Earlier, Caldwell (1986) described instances in which Keralans held health workers accountable through strong-armed means: "Doctors and others who provide village services (for instance, bus drivers plying regular routes) know stories of their fellows who were treated violently or hurt in protests about their having failed their duty" (p. 199). He quotes from a colleague working in West Bengal, where success was achieved "because [the state's communist government] appointed cadres at every health center to report on doctors or nurses who did not put all their time and effort into their services or who discriminated between patients" (p. 203).

A second example comes from an assessment of a major health intervention in Ceara State, northeastern Brazil, that contributed to a 36 percent fall in infant mortality in only a few years. Tendler and Freedheim (1994:1777–78) identify three primary reasons behind this success. First, the state used a merit hiring system and a large advertising campaign to create "a sense of 'mission' around the program and remarkable respect for its workers in the communities in which they served." Second, flexibility in job descriptions allowed workers to take on tasks that, although "sometimes viewed as distractions by experts, formed the basis for relations of trust between workers and citizens." Third, job candidates who were rejected were educated about what to expect from workers, supervisors, and elected officials, turning them into "informed public monitors of a new program in which the potential for

abuse was high." For example, job applicants were told, "Those of you who are not selected must make sure that those who are chosen abide by the rules. . . . If these rules are breached we want to hear about it. . . . [W]e are keeping all the applications, just in case any of those we hire do not perform well."

Motivating public (and private) sector workers to deliver high-quality services is a long-standing issue (and not just in the health sector), especially in situations where monitoring performance is difficult or costly. In the examples here, such motivation seems to have been achieved through direct monitoring with the threat of job loss, community monitoring with various threats, and community oversight and participation with the threat of job (and prestige) loss.

Public Sector Provision and the Total Use of Services

Why does the evidence show that even large-scale provision of some types of primary health care services has little or no impact on health status? One reason is that extending publicly funded health care could merely crowd out the consumption of equally effective private services. Even if the government were to deliver health services effectively, the health impact would depend not on the total use of public services but rather on how public provision affected total use of all services.

Although one might think at first blush that overall service use would increase by the same amount as the increase in government service capacity, the net effect of increasing public supply is always less than one for one. Only if there were no private sector or if there were no substitution between public and private services would total service increase by the same amount as the increase in government service. The degree to which public spending affects overall health services depends on four factors.[6]

- *The smaller the overall elasticity of demand for health services, the smaller the impact of public spending.* Low elasticity of demand (demand that is unresponsive to price and distance) may well be the crux of the matter. Economists tend to shy away from discussions of "need" because it is not directly observable and is an emotionally charged term. But if an individual "needs" something, she or he is likely to forgo other consumption in order to satisfy that need. If income falls or prices rise, consumption of such a good will be protected, and the adjustment will be greater with regard to other uses of money and time. Because it is assumed that individuals place a high value on necessary health services (or at the very least those that alleviate unpleasant symptoms), the demand for them would be fairly insensitive to circumstances such as price or availability. Therefore, for those health services that are a matter of life or death, low price elasticities of demand should be expected. To the extent that severe symptoms are associated with health problems that are best treated in a clinic (which is not always the case),

these are the problems that are likely to be treated more or less successfully, regardless of a public presence. Makinen and Raney (1994:17) summarize five studies of medical care demand and find consistently that "price influences choice among providers and does so more strongly than its influence on whether or not to use services at all."

Indeed, in China (Cretin and others 1990), Indonesia (Gertler and Molyneaux 1995), and the United States (Manning and others 1987; Newhouse 1995), studies confirm that the demand for treatment is less elastic for serious conditions than for less serious conditions. Market failures associated with the treatment of serious health problems are likely to manifest themselves in ways other than bad health. For example, people may be overcharged, service may be inconvenient, or assets may be liquidated at inopportune moments (to get ready cash). All of these may result in inefficiency. But the impact on health, especially mortality, will be small.

The degree to which people take care of their health problems depends on the time elapsing between the onset of symptoms and treatment. In emergency situations, to take an extreme case, the effectiveness of treatment is tightly time-bound, allowing no opportunity to shop around. The demand for treatment will be inelastic to its effective price. However, other dimensions of access that act as barriers to emergency care, such as emergency transport or credit, might be amenable to government interventions.

Health problems that can be treated if detected early but that have symptoms which are relatively minor at early stages may generate a demand for examinations that is highly elastic with respect to price or accessibility. In such cases a different dynamic provides another way in which government subsidies may well increase the socially optimal use of services. Screening for hypertension and some cancers falls into this category. This line of reasoning relies on subtle interactions of the severity of symptoms, the effectiveness of treatment, and, most important, the responsiveness to price and distance of the demand for care. In general, however, one would expect more serious illnesses to be associated with more inelastic demand for public services and, therefore, less potential influence of government policy on health status.

- *The larger the private sector, the smaller the impact of public spending.* According to the *World Development Indicators* (World Bank 1999), between 1990 and 1997 private (nongovernment) spending made up the largest percentage of health spending in nearly all poor countries, accounting for 59 percent in low-income countries (those with gross national product per capita of $785 or less in 1997) and almost 75 percent in South Asia (table 2).

Direct evidence on visits and treatments is hard to come by. A study in five Indian states, summarized in table 3, shows that in rural areas 82 percent of treated illness episodes were treated by private providers and that expenditures

Table 2. *Share of Total Health Expenditures That Is Private*

Category	*Average (percent)*	*Number of countries*
All countries	50.9	123
Income group		
Low income	58.7	34
Lower middle income	40.2	34
Higher middle income	47.9	24
High income, non-OECD	49.1	8
High income, OECD	33.3	23
Regions		
East Asia and Pacific	47.2	14
Latin America and the Caribbean	58.5	35
Middle East and North Africa	48.8	12
South Asia	74.9	5
Sub-Saharan Africa	51.7	22
Other	28.5	35

Notes: Values are population-weighted means of country averages for the 1990–97 period. OECD, Organisation for Economic Co-operation and Development.

Source: Derived from the World Development Indicators 1999 database.

for nonhospital treatment were nearly two-thirds (65 percent) of all out-of-pocket expenditures (World Bank 1995a). Note that private providers receive a larger part of these nonhospital expenditures (86 percent) than of expenditures as a whole because people can pay out of pocket for relatively minor illnesses but are less able to cover hospital expenses.

- *The greater the extent to which people see the private sector as a substitute for the public sector, the smaller the impact of public spending.* In addition to the size of the private sector, there is mounting evidence on the substitutability between public and private health care providers. Users may substitute private for public providers on the basis of price, location of facility, or quality of services. Several studies on the effect of increases in fees at (or distance to) public clinics look at the probability that a sick individual will attend a public versus a private clinic (table 4). Although results are not the same across all countries, both sets of characteristics have an effect on demand for care from private providers. The "cross-price" effect—for example, the impact of a change in fees at private clinics on the use of public clinics—is not negligible, especially if compared with "own-price" effects. For example, in Nigeria 100 percent of those who are deterred by higher prices in the public sector sought care from the private sector; in Ghana 60 percent of those deterred went to private providers; in El Salvador the fraction was 50 percent. Clearly, there is a high degree of substitutability between the two types of providers. Similarly, construction of new public facilities closer to consumers (reducing travel distance) can also displace demand from private

Table 3. *Demand for Private Sector Health Care in Five Indian States*

Area	*Gujarat*	*Maharashtra*	*Tamil Nadu*	*Uttar Pradesh*	*West Bengal*	*Weighted average*
Rural						
Percentage of treated illness episodes going to private providers	69	78	71	91	83	82
Percentage of out-of-pocket spending going to nonhospital treatment	62	64	74	59	74	65
Percentage of out-of-pocket spending going to private nonhospital treatment	48	55	68	51	65	56
Percentage of out-of-pocket spending on nonhospital treatment going to private treatment	77	86	92	86	88	86
Urban						
Percentage of treated illness episodes going to private providers	82	75	69	85	78	79
Percentage of out-of-pocket spending going to nonhospital treatment	61	60	69	57	63	61
Percentage of out-of-pocket spending going to private nonhospital treatment	53	51	62	46	54	52
Percentage of out-of-pocket spending on nonhospital treatment going to private treatment	87	85	90	81	86	85

Source: Adapted from World Bank (1995a).

providers in the vicinity of the new facility. This effect is estimated to be substantial in Ghana and Kenya.

Another factor influencing the degree of substitutability between public and private providers is the absolute (expected) expense of the treatment sought. As discussed above, patients bypass both public and private facilities depending on their circumstances. Without insurance, people cannot afford to pay for expensive treatments when hospital care is essential; there is no plausible alternative to a public facility. For relatively cheap treatments, a private sector can thrive even without health insurance, and it is for such treatments that substitution between the sectors will be high.

- *The larger the private sector response to public intervention, the smaller the impact of public spending.* Changes in the price or availability of government interventions may induce a private supply response that can mitigate any actual impact on health outcomes. A case-control experiment in Indonesia showed that increasing user fees in public facilities caused a decline in use of those facilities and an expansion of the private health care sector; as a result, the effect of a user fee

Table 4. *Studies on the Probability of Provider Chosen: Percentage Deterred from Public Facilities Who Go to Private Facilities*

Sample	*Change*	*Percent*
Increase in price		
Benin (rural)	Community health center fees	61
Bolivia (urban)	Fees in Ministry of Public Health facilities, effect on adults 16 and over	33
	Fees in Ministry of Public Health facilities, effect on children 15 and under	0
	Fees in Ministry of Public Health facilities, effect on children 4 and under	0
El Salvador (urban)	Fees in Ministry of Health facilities, effect on males	56
	Fees in Ministry of Health facilities, effect on females	50
Ghana	Public fees	60
Kenya (rural)	Fees in government facilities from 0 to 10 Kenyan shillings, effect on adults 15 and over	37
Nigeria	Public prices, effect on adults 16 and over	100
Pakistan (urban)	Government clinic price, effect on children 5 and under	71
Increase in distance or time		
Bolivia (urban)	Waiting time in Ministry of Public Health facilities, effect on adults 16 and over	100
	Travel time to Ministry of Public Health facilities, effect on children 15 and under	20
	Waiting time in Ministry of Public Health facilities, effect on children 4 and under	20
Ghana	Distance to nearest public facility	50
Kenya (rural)	Distance to government facilities	44

Note: See Filmer, Pritchett and Hammer (1998) for more details on the results in the underlying studies

Sources: Derived from the following sources: Benin—Bolduc, Lacroix, and Muller (1996); Bolivia—Ii (1996); El Salvador—Bitran and McInnes (1993); Ghana—Lavy and Germain (1994); Kenya—Mwabu, Ainsworth, and Nyamete (1993); Nigeria—Akin, Guilkey, and Denton (1995); Pakistan—Alderman and Gertler (1989).

increase on overall health outcomes was very small (Gertler and Molyneaux 1995). In Tanzania, after the provision of health care was liberalized in 1991, the number of private sector providers skyrocketed; the number of nonprofit units operated by "approved voluntary organizations" increased from 697 to 780, and the number of private, for-profit facilities increased from 41 to 1,340 (Munishi 1997). The private sector expansion was dominated numerically by lower-level facilities such as dispensaries (from 36 to 1,313), but the number of private, for-profit hospitals also increased, from 4 to 20, while the number of government hospitals remained constant at 77.

A review of public spending in 13 regions in the Philippines (World Bank 1995b) from 1983 to 1990 found that public spending on health improved infant mortality in regions with lower incomes but had virtually no effect in richer regions. This result is consistent with substitutability of the private and public sectors in higher-income regions, with a resulting highly elastic private supply (and small impact on outcomes) and a less elastic supply response in poorer regions, with a substantial impact of public spending (Hammer 1997).

Although analyzing the potential impact of public sector interventions on the private sector might seem an obvious undertaking, it often has been ignored. A recent internal review of the World Bank—an institution with a large number of economists—found that of 217 appraisals of Bank health projects since 1970, *only 4* (all of these since 1995) had developed hypotheses regarding (or even considering) potential crowding-out (or crowding-in) as a result of the public supply of health inputs (World Bank 1998b).

Conclusion

Why is the evidence so thin for the seemingly plausible argument that primary health care is the best way to improve health in developing countries? We have argued that there are two weak links in the chain between providing public money for primary health care activities and achieving better health outcomes. The first is that governments often find it difficult to translate public spending into effective services, and the second is that of all the types of health services that the public sector might provide, primary care is the one that the private sector is likely to move away from as a result of any increase in public supply. Neither of these points is usually considered when discussing the applicability of a primary health care paradigm to a particular country.

What can be done? An alternative way of looking at policy in the health sector focuses on a careful analysis of the characteristic market failures in a country's health sector, the degree to which policy benefits reach the poor, and the government's ability to provide various kinds of services. Improvement of the quality of government health services depends on pressures outside the bureaucracy—citizen voice and consumer choice. From this perspective, clear recommendations can follow. First, public health activities should emphasize control of infectious diseases, largely through environmental changes. Second, public health programs should include demand-side mechanisms to improve routine clinical care. And third, policymakers should reconsider the role of hospital services as a practical means for governments to address inadequacies in insurance markets. The details of these arguments are available in a companion paper, "A Prescription for Health Policy in Poor Countries" (available from the authors).

Notes

Deon Filmer is an economist with the World Bank's Development Research Group, Jeffrey S. Hammer is lead economist with the Development Research Group, and Lant H. Pritchett is a lecturer at the John F. Kennedy School of Government, Harvard University, on leave from the World Bank. This paper grew out of an earlier collaboration between the authors and Maureen Lewis and Samuel Lieberman. The authors thank Philip Musgrove, Martin Ravallion, and Susan Stout for helpful discussions and comments.

1. This statistic is based on a comparison of average mortality rates for children under five in low- and middle-income countries with rates in high-income countries (World Bank 1999; see also Gwatkin 1980).

2. An entirely different viewpoint, also common in the literature, is that medical care of any sort—public or private—has had little, if any, impact on overall health measures. In this view, the large declines in mortality in the West preceded the development of medical treatments effective in combating its major causes. Better nutrition and housing conditions resulting from higher income as well as better hygienic practices are more likely responsible. We do not dispute this interpretation and note that it argues even more strongly for getting the first link right—making sure the composition of spending is geared toward basic hygiene—safe water, sanitation, and education.

3. Along with nearly all of the rest of the literature, we assume a fixed budget and avoid the problem of maximizing welfare, as that would endogenize the budget and require valuing health versus nonhealth goods.

4. In addition, the huge effort to create measures of disability-adjusted life years (DALYs) has led to some additional information on morbidity (Murray 1994). However, the correlation between DALYs lost and life expectancy or infant mortality is a high 0.93 across the eight regions for which DALYs have been calculated (World Bank 1993).

5. In that analysis, two-stage least-squares estimation was used to address the potential problems of measurement error and reverse causation. The estimates will be biased toward zero if these variables are measured with error (this is likely because a value for 1985 is sometimes used in place of one for 1990). In addition, the estimates will be inconsistent if there is reverse causation—if, for example, high mortality rates induce the government to spend more on access to local health services.

6. In other language, if D is total demand, S_b is public supply, ε_D is the elasticity of demand, ε_S is the elasticity of supply, and D_v is the demand for private services, then, allowing public and private services to be imperfect substitutes, with the elasticity of demand for private services with respect to public services $\varepsilon_{vb} < 0$, leads to $(dD/dS_b) = 1 + [(\varepsilon_S \times \varepsilon_{vb})/(\varepsilon_S - \varepsilon_D)] \times D_v/S_b$.

References

The word "processed" refers to informally reproduced works that may not be commonly available through library systems.

Agrawal, Amar Nath, Hari Om Varma, and R. C. Gupta, eds. 1996. *India: Economic Information Year-Book.* New Delhi: National.

Akin, John S., and Paul Hutchinson. 1999. "Health-Care Facility Choice and the Phenomenon of Bypassing." *Health Policy and Planning* 14(2):135–51.

Akin, John S., David K. Guilkey, and E. Hazel Denton. 1995. "Quality of Services and Demand for Health Care in Nigeria: A Multinomial Probit Estimation." *Social Science and Medicine* 40(11): 1527–37.

Alderman, Harold, and Paul Gertler. 1989. *The Substitutability of Public and Private Health Care for the Treatment of Children in Pakistan.* Living Standards Measurement Study Working Paper 57. Washington, D.C.: World Bank.

Alderman, Harold, and Victor Lavy. 1996. "Household Responses to Public Health Services: Cost and Quality Tradeoffs." *The World Bank Research Observer* 11(1):3–22.

Anand, Sudhir, and Martin Ravallion. 1993. "Human Development in Poor Countries: On the Role of Private Incomes and Public Services." *Journal of Economic Perspectives* 7(1):133–50.

Benefo, Kofi, and T. Paul Schultz. 1996. "Fertility and Child Mortality in Côte d'Ivoire and Ghana." *The World Bank Economic Review* 10(1):123–58.

Bidani, Benn, and Martin Ravallion. 1997. "Decomposing Social Indicators Using Distributional Data." *Journal of Econometrics* 77:125–39.

Bitran, Ricardo A., and D. Keith McInnes. 1993. *The Demand for Health Care in Latin America: Lessons from the Dominican Republic and El Salvador.* Economic Development Institute Seminar Paper 46. Washington, D.C.: World Bank.

Bolduc, Denis, Guy Lacroix, and Christophe Muller. 1996. "The Choice of Medical Providers in Rural Benin: A Comparison of Discrete Choice Models." *Journal of Health Economics* 15:477–98.

Caldwell, John C. 1986. "Routes to Low Mortality in Poor Countries." *Population and Development Review* 12(2):171–220.

Cretin, S., N. Duan, A. Williams, X. Gu, and Y. Shi. 1990. "Modeling the Effect of Insurance on Health Expenditures in the People's Republic of China." *Health Services Research* 25(4):667–85.

DaVanzo, Julie. 1984. "A Household Survey of Child Mortality Determinants in Malaysia." *Population and Development Review* 10(Supplement):307–23.

Decosas, Josef. 1990. "Planning for Primary Health Care: The Case of the Sierra Leone National Action Plan." *International Journal of Health Services* 20(1):167–77.

Filmer, Deon, and Lant Pritchett. 1999. "The Impact of Public Spending on Health: Does Money Matter?" *Social Science and Medicine* 49(1):1309–23.

Filmer, Deon, Lant Pritchett, and Jeffrey Hammer. 1998. "Health Policy in Poor Countries: Weak Links in the Chain?" Policy Research Working Paper 1874. World Bank, Development Research Group, Poverty and Human Resources, Washington, D.C. Processed.

Foster, S. D. 1990. "Improving the Supply and Use of Essential Drugs in Sub-Saharan Africa." Policy Research Working Paper 456. World Bank, Population and Human Resources Department, Washington, D.C. Processed.

Fox, Tamara. 1995. "Primary Health Care: Notes on the Literature and Annotated Bibliography." Paper prepared for the World Bank, Latin America and Caribbean Human Development Group, Washington, D.C. Processed.

Frankenberg, Elizabeth. 1995. "The Effects of Access to Health Care on Infant Mortality in Indonesia." *Health Transition Review* 5:143–63.

Gertler, Paul, and John Molyneaux. 1995. "Pricing Public Health Services: Lessons from a Social Experiment in Indonesia." RAND Corporation. Santa Monica, Calif. Processed.

Griffin, Charles C. 1992. *Health Care in Asia: A Comparative Study of Cost and Financing.* Washington, D.C.: World Bank.

Gwatkin, Davidson R. 1980. "How Many Die? A Set of Demographic Estimates of the Annual Number of Infant and Child Deaths in the World." *American Journal of Public Health* 70:1286–89.

Hammer, Jeffrey S. 1997. "Economic Analysis for Health Projects." *The World Bank Research Observer* 12(1):47–71.

Hammer, Jeffrey S., Ijaz Nabi, and James Cercone. 1995. "Distributional Effects of Social Sector Expenditures in Malaysia, 1974 to 1989." In Dominique van de Walle and Kimberly Nead, eds., *Public Spending and the Poor: Theory and Evidence.* Baltimore, Md.: Johns Hopkins University Press.

Heller, Patrick. 1996. "Social Capital as a Product of Class Mobilization and State Intervention: Industrial Workers in Kerala, India." *World Development* 24(6):1055–71.

Hossain, Shaikh. 1989. "Effect of Public Programs on Family Size and Child Education and Health." *Journal of Development Economics* 30(1):144–58.

Ii, Masako. 1996. *The Demand for Medical Care: Evidence from Urban Areas in Bolivia.* Living Standards Measurement Study Working Paper 123. Washington, D.C.: World Bank.

International Institute for Population Studies. 1995. *National Family Health Survey (MCH and Family Planning) India 1992–93.* Bombay.

Koenig, Michael A., Vincent Fauveau, and Bogodan Wojtyniak. 1991. "Mortality Reductions from Health Interventions: The Case of Immunization in Bangladesh." *Population and Development Review* 17:87–104.

Kunstadter, Peter. 1985. "Remarks." In Scott B. Halstead, Julia A. Walsh, and Kenneth S. Warren, eds., *Good Health at Low Cost.* New York: Rockefeller Foundation.

Lavy, Victor, and Jean-Marc Germain. 1994. *Quality and Cost in Health Care Choice in Developing Countries.* Living Standards Measurement Study Working Paper 105. Washington, D.C.: World Bank.

Lavy, Victor, John Strauss, Duncan Thomas, and Philippe de Vreyer. 1996. "Quality of Health Care, Survival, and Health Outcomes in Ghana." *Journal of Health Economics* 15:333–57.

Lewis, Maureen, Gunnar S. Eskeland, and Ximena Traa-Valerezo. 1999. "Challenging El Salvador's Rural Health Care Strategy." Policy Research Working Paper 2164. World Bank, Washington, D.C. Processed.

Lewis, Maureen A., Gerald M. La Forgia, and Margaret B. Sulvetta. 1996. "Measuring Public Hospital Costs: Empirical Evidence from the Dominican Republic." *Social Science and Medicine* 43(2):221–34.

Liu Yuanli, William C. Hsiao, Qing Li, Xingzhu Liu, and Minghui Ren. 1995. "Transformation of China's Rural Health Care Financing." *Social Science and Medicine* 41(8):1085–93.

Macro International. 1997. *Brasil Pesquisa Nacional sobre Demographia e Saude, 1996.* Demographic and Health Survey Program. Calverton, Md.

Makinen, Marty, and Laura Raney. 1994. "Role and Desirability of User Charges for Health Services." Draft Working Paper. Abt Associates, Bethesda, Md. Processed.

Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, Arleen Leibowitz, and M. Susan Marquis. 1987. "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment." *American Economic Review* 77(3):251–77.

Menken, Jane, and James Phillips. 1990. "Population Change in a Rural Area of Bangladesh, 1967–87." *The Annals.* 510:87–101.

Milgrom, Paul, and John Roberts. 1992. *Economics, Organization, and Management.* Englewood Cliffs, N.J.: Prentice-Hall.

Mtemeli, D. N. 1994. "An Investigation into Why Women in Mutasa District Prefer to Deliver at Home and Not in Health Institutions." Provincial Medical Direction, Manicaland, Zimbabwe. Processed.

Muhuri, Pradip K., and Samuel H. Preston. 1991. "Effects of Family Composition on Mortality Differentials by Sex among Children in Matlab, Bangladesh." *Population and Development Review* 17:415–34.

Munishi, Gaspar K. 1997. "Private Health Sector Growth following Liberalization in Tanzania: Some Policy Considerations." International Health Policy Program Working Paper. Washington, D.C. Processed.

Murray, Christopher. 1994. "Quantifying the Burden of Disease: The Technical Basis for Disability-Adjusted Life Years." *Bulletin of the World Health Organization* 72(3):429–45.

Musgrove, Philip. 1996. *Public and Private Roles in Health: Theory and Financing Patterns.* World Bank Discussion Paper 339. Washington, D.C.

Mwabu, Germano, Martha Ainsworth, and Andrew Nyamete. 1993. "Quality of Medical Care and Choice of Medical Treatment in Kenya: An Empirical Analysis." *Journal of Human Resources* 28(4):838–62.

Newhouse, Joseph P. 1995. *Free for All.* Cambridge, Mass.: Harvard University Press.

Paes de Barros, Ricardo, D. R. Da Costa, and Rosane Mendoça. 1998. "Determinants of Infant Mortality in Brazil: Evidence from Census Data." Working Paper. Instituto de Pesquisa Económica Aplicada, Brasilia. Processed.

Panis, Constantijin W. A., and Lee A. Lillard. 1994. "Health Inputs and Child Mortality in Malaysia." *Journal of Health Economics* 13(4):455–89.

PIEDR (Pakistan Institute for Environment-Development Action) and Project Management Team. 1994. "The State of the Public Sector Primary Health Care Services, District Sheikhupura, Punjab, Pakistan." Bamako Initiative Technical Report Series 31. Bamako Initiative Management Unit, United Nations Children's Fund, New York. Processed.

Pitt, Mark, Mark Rosenzweig, and Donna Gibbons. 1993. "The Determinants and Consequences of the Placement of Government Programs in Indonesia." *The World Bank Economic Review* 7(3):319–48.

Preston, Samuel H. 1980. "Causes and Consequences of Mortality in Less Developed Countries during the Twentieth Century." In Richard Easterlin, ed., *Population and Economic Change in Developing Countries.* Chicago, Ill.: University of Chicago Press for the National Bureau of Economic Research.

———. 1986. "Mortality and Development Revisited." *Population Bulletin of the United Nations.* No. 18-1985. Department of International Economic and Social Affairs, United Nations, New York.

Ratzan, Scott C., Gary L. Filerman, and John W. LeSar. 2000. "Attaining Global Health: Challenges and Opportunities." *Population Bulletin* 55(1):1–48.

Rosenzweig, Mark, and T. P. Schultz. 1982. "Child Mortality and Fertility in Colombia: Individual and Community Effects." *Health Policy and Education* 2:305–48.

Rosenzweig, Mark, and Kenneth Wolpin. 1982. "Government Interventions and Household Behavior in a Developing Country: Anticipating the Unanticipated Consequences of Social Programs." *Journal of Development Economics* 10:209–25.

Samrasinghe, Daya, and John Akin. 1994. "Health Strategy and Financing Study: Final Report." Prepared for Government of Sri Lanka, Ministry of Health and Women's Affairs, Colombo. Processed.

Sastry, Narayan. 1995. "Community Characteristics, Individual and Household Attributes, and Child Survival in Brazil." RAND Labor and Population Program Working Paper Series 95-13. RAND Corporation, Santa Monica, Calif. Processed.

Taylor, Carl E., and R. D. Singh. n.d. "The Narangwal Population Study: Integrated Health and Family Planning Services." Rural Health Research Center, Narangwal, India. Processed.

Tendler, Judith, and Sara Freedheim. 1994. "Trust in a Rent Seeking World: Health and Government Transformed in Northeastern Brazil." *World Development* 22(2):1771–91.

Thomas, Duncan, Victor Lavy, and John Strauss. 1996. "Public Policy and Anthropometric Outcomes in Côte d'Ivoire." *Journal of Public Economics* 61:155–92.

WHO (World Health Organization). 1988. *Alma-Ata to the Year 2000: Reflections at the Midpoint.* Geneva.

World Bank. 1993. *World Development Report 1993: Investing in Health.* New York: Oxford University Press.

———. 1994a. *Better Health in Africa: Lessons and Experiences Learned.* Washington, D.C.

———. 1994b. "Indonesia: Indonesia's Workforce: Issues and Options." Report 1283S. Population and Human Resource Division, East Asia and Pacific Regional Office, Washington, D.C. Processed.

———. 1995a. "India: Policy and Finance Strategies for Strengthening Primary Health Care Services." Report 13042-IN. Washington, D.C. Processed.

———. 1995b. "Philippines: Public Expenditure Review." East Asia Department I, Washington, D.C. Processed.

———. 1998a. *Reducing Poverty in India: Options for More Effective Public Services.* World Bank Country Study. W ashington, D.C.

———. 1998b. "Lessons from Experience in Health, Nutrition, and Population." Sector and Thematic Evaluations Group, Operations Evaluation Department, Washington, D.C. Processed.

———. 1999. *World Development Indicators.* Washington, D.C.

Social Capital: Implications for Development Theory, Research, and Policy

Michael Woolcock • Deepa Narayan

In the 1990s the concept of social capital—defined here as the norms and networks that enable people to act collectively—enjoyed a remarkable rise to prominence across all the social science disciplines. The authors trace the evolution of social capital research as it pertains to economic development and identify four distinct approaches the research has taken: communitarian, networks, institutional, and synergy. The evidence suggests that of the four, the synergy view, with its emphasis on incorporating different levels and dimensions of social capital and its recognition of the positive and negative outcomes that social capital can generate, has the greatest empirical support and lends itself best to comprehensive and coherent policy prescriptions. The authors argue that a significant virtue of the idea of and discourse on social capital is that it helps to bridge orthodox divides among scholars, practitioners, and policymakers.

What is social capital? How does it affect economic development? What are the implications for theory, research, and policy? These questions lie at the heart of recent attempts to make sense of the burgeoning literature on social capital and to ascertain its relationship to economic development. In this article we endeavor to answer each of these questions; in so doing, we provide an overview of the scholarship on social capital for those unfamiliar with the term as well as a sense of coherence and direction to those embarking on new empirical research and policy analysis in this rich field.

What Is Social Capital?

"It's not what you know, it's *who* you know." This common aphorism sums up much of the conventional wisdom regarding social capital. It is wisdom born of

The World Bank Research Observer, vol. 15, no. 2 (August 2000), pp. 225–49.

experience—that gaining membership to exclusive clubs requires inside contacts, that close competitions for jobs and contracts are usually won by those with friends in high places. When people fall on hard times, they know it is their friends and family who constitute the final safety net. Conscientious parents devote hours to the school board and to helping their children with homework, only too aware that a child's intelligence and motivation are not enough to ensure a bright future. Some of our happiest and most rewarding hours are spent talking with neighbors, sharing meals with friends, participating in religious gatherings, and volunteering for community projects.

Intuitively, then, the basic idea of social capital is that a person's family, friends, and associates constitute an important asset, one that can be called on in a crisis, enjoyed for its own sake, and leveraged for material gain. What is true for individuals, moreover, also holds for groups. Those communities endowed with a diverse stock of social networks and civic associations are in a stronger position to confront poverty and vulnerability (Moser 1996; Narayan 1995), resolve disputes (Schafft 1998; Varshney 2000), and take advantage of new opportunities (Isham 1999). Conversely, the *absence* of social ties can have an equally important impact. Office workers, for example, fear being left out of the loop on important decisions; ambitious professionals recognize that getting ahead in a new venture typically requires an active commitment to networking. A defining feature of being poor, moreover, is that one is not a member of—or may even be actively excluded from—certain social networks and institutions that could be used to secure good jobs and decent housing (Wilson 1987, 1996).

Intuition and everyday language also recognize an additional feature of social capital: that it has costs as well as benefits, that social ties can be a liability as well as an asset. Most parents, for example, worry that their teenage children will fall in with the wrong crowd and that peer pressure and a strong desire for acceptance will induce them to take up harmful habits. Even close family members can overstay their welcome. At the institutional level, many countries and organizations have nepotism laws, in explicit recognition that personal connections can be used to discriminate unfairly, distort, and corrupt. Everyday language and life experience, in short, teach that the social ties individuals have can be both a blessing and a blight, while those they do *not* have can deny them access to key resources. These features of social capital are well documented by the empirical evidence and have important implications for economic development and poverty reduction.

These examples suggest a more formal definition: social capital refers to the norms and networks that enable people to act collectively. This simple definition serves a number of purposes. First, it focuses on the sources, rather than the consequences, of social capital (Portes 1998) while recognizing that important features of social capital, such as trust and reciprocity, are developed in an iterative process. Second, this definition permits the incorporation of different dimensions of social capital and

recognizes that communities can have access to more or less of them. The poor, for example, may have a close-knit and intensive stock of "bonding" social capital that they can leverage to "get by" (Briggs 1998; Holzmann and Jorgensen 1999), but they lack the more diffuse and extensive "bridging" social capital deployed by the nonpoor to "get ahead" (Barr 1998; Kozel and Parker 2000; Narayan 1999). Accordingly, such an approach allows the argument that it is different *combinations* of bonding and bridging social capital that are responsible for the range of outcomes observed above and incorporates a dynamic component in which optimal combinations of these dimensions change over time. Third, while this definition presents the community (rather than individuals, households, or the state) as the primary unit of analysis, it recognizes that individuals and households (as members of a given community) can nonetheless appropriate social capital and that the way communities themselves are structured turns in large part on their relationship with the state. Weak, hostile, or indifferent governments have a profoundly different effect on community life and development projects, for example, than do governments that respect civil liberties, uphold the rule of law, honor contracts, and resist corruption (Isham and Kaufmann 1999).

This conceptualization of the role of social relationships in development represents an important departure from earlier theoretical approaches and therefore has important implications for contemporary development research and policy. Until the 1990s the major theories of development held rather narrow, even contradictory, views about the role of social relationships in economic development and offered few constructive policy recommendations. In the 1950s and 1960s, for example, traditional social relationships and ways of life were viewed as impediments to development. When modernization theorists explained "the absence or failure of capitalism," Moore (1997:289) correctly notes, "the focus [was] on social relations as obstacles." As an influential United Nations (1951) document of the time put it, for development to proceed, "ancient philosophies have to be scrapped; old social institutions have to disintegrate; bonds of caste, creed and race have to burst; and large numbers of persons who cannot keep up with progress have to have their expectations of a comfortable life frustrated" (cited in Escobar 1995:3).

This view gave way in the 1970s to the arguments of dependency and world-systems theorists, who held that social relations among corporate and political elites were a primary mechanism of capitalist exploitation. The social characteristics of poor countries and communities were defined almost exclusively in terms of their relation to the means of production and the inherent antipathy between the interests of capital and labor. Little mention was made of the possibility (or desirability) of mutually beneficial relationships between workers and owners, of the tremendous variation in the degree of success recorded by developing countries, or of political strategies—other than revolution—by which the poor could improve their lot. At the same time, communitarian perspectives stressed the inherent beneficence and

self-sufficiency of local communities but underestimated the negative aspects of communal obligations, overestimated the virtues of isolationism and self-sufficiency, and neglected the importance of social relations in constructing effective and accountable formal institutions. For their part, neoclassical and public choice theorists—whose voices were the most influential in the 1980s and early 1990s—assigned no distinctive properties to social relations. These perspectives, which focused on the strategic choices of rational individuals interacting under various time, budgetary, and legal constraints, held that groups (including firms) existed primarily to lower the transaction costs of exchange; given undistorted market signals, the optimal size and combination of groups would duly emerge.

The major development theories, then, construed social relations as singularly burdensome, exploitative, liberating, or irrelevant. Reality, unfortunately, does not conform so neatly to these descriptions and their corresponding policy prescriptions. Events in the post–cold war era—from ethnic violence and civil war to financial crises and the acknowledgement of widespread corruption—demand a more sophisticated appraisal of the vices, virtues, and vicissitudes of the social dimension as it pertains to the wealth and poverty of nations (Woolcock forthcoming). The literature on social capital, in its broadest sense, represents a first approximation to the answer to this challenge. It is a literature to which all the social science disciplines have contributed, and it is beginning to generate a remarkable consensus regarding the role and importance of institutions and communities in development. Indeed, one of the primary benefits of the idea of social capital is that it allows scholars, policymakers, and practitioners from different disciplines to enjoy an unprecedented level of cooperation and dialogue (Brown 1998; Brown and Ashman 1996).

Four Perspectives on Social Capital and Economic Development

The letter and spirit of social capital have a long intellectual history in the social sciences (Platteau 1994; Woolcock 1998), but the sense in which the term is used today dates back more than 80 years to the writings of Lyda J. Hanifan, then the superintendent of schools in West Virginia. Explaining the importance of community participation in enhancing school performance, Hanifan (1916:130) invoked the concept of social capital, describing it as

> those tangible substances [that] count for most in the daily lives of people: namely good will, fellowship, sympathy, and social intercourse among the individuals and families who make up a social unit. . . . If [an individual comes] into contact with his neighbor, and they with other neighbors, there will be an accumulation of social capital, which may immediately satisfy his

> social needs and which may bear a social potentiality sufficient to the substantial improvement of living conditions in the whole community.

After Hanifan the idea of social capital disappeared for several decades but was reinvented in the 1950s by a team of Canadian urban sociologists (Seely, Sim, and Loosely 1956), in the 1960s by an exchange theorist (Homans 1961) and an urban scholar (Jacobs 1961), and in the 1970s by an economist (Loury 1977). None of these writers, interestingly, cited earlier work on the subject, but all used the same umbrella term to encapsulate the vitality and significance of community ties. The seminal research by Coleman (1987, 1988, 1990) on education and by Putnam (1993, 1995) on civic participation and institutional performance, however, has provided the inspiration for most of the current work, which has since coalesced around studies in nine primary fields: families and youth behavior; schooling and education; community life (virtual and civic); work and organizations; democracy and governance; collective action; public health and environment; crime and violence; and economic development.[1]

In this paper we are concerned with this final category and related work in political economy and new institutional economics. Research on social capital and economic development can be categorized into four distinct perspectives: the communitarian view, the networks view, the institutional view, and the synergy view.

The Communitarian View

The communitarian perspective equates social capital with such local organizations as clubs, associations, and civic groups. Communitarians, who look at the number and density of these groups in a given community, hold that social capital is inherently good, that more is better, and that its presence always has a positive effect on a community's welfare. This perspective has made important contributions to analyses of poverty by stressing the centrality of social ties in helping the poor manage risk and vulnerability. As Dordick (1997) notes, the poor have "something left to lose"—each other.

In their celebration of community and civil society, however, many enthusiasts of this view of social capital have ignored its important downside (Portes and Landolt 1996). For example, where communities or networks are isolated, parochial, or working at cross-purposes to society's collective interests (in ghettos, gangs, drug cartels, and so on), productive social capital is replaced by what Rubio (1997)—in discussing Colombia—calls *perverse* social capital, which greatly hinders development. Many benefits certainly are associated with being a member of a highly integrated community, but there are also significant costs, and for some, the costs may greatly outweigh the benefits. Consider, for instance, the bright girls who are taken out of village schools in India because of community expectations. The social networks underly-

ing organized crime syndicates in Latin America and Russia may generate large negative externalities for society in the form of lost lives, wasted resources, and pervasive uncertainty. The communitarian perspective also implicitly assumes that communities are homogenous entities that automatically include and benefit all members. But the extensive literature on caste inequality, ethnic exclusion, and gender discrimination—the bleak outcomes often produced and maintained by community pressures—suggests otherwise (Narayan and Shah 1999).

Evidence from the developing world demonstrates why merely having high levels of social solidarity or informal groups does not necessarily lead to economic prosperity. In Kenya a participatory poverty assessment recorded more than 200,000 community groups active in rural areas, but most were unconnected to outside resources and were unable to improve the lot of the poor (Narayan and Nyamwaya 1996). A World Bank (1989) report on Rwanda cited more than 3,000 registered cooperatives and farmers groups and an estimated 30,000 informal groups, yet these groups were unable to prevent one of history's most gruesome civil wars. In many Latin American countries, indigenous groups are often marked by high levels of social solidarity, but they remain excluded economically because they lack the resources and access to power that are necessary to shift the rules of the game in their favor (Narayan 1999). This is also the case in Haiti, where social capital, "rich at the local level," is employed by peasant groups to "meet labor requirements, gain access to land, protect clientship in the marketplace, promote mutual aid, assure protection from state authorities, and generally manage risk." Even so, these groups cannot overcome the crippling effects of colonialism, corruption, "geographical isolation, political exclusion, and social polarization" (all quotations from White and Smucker 1998:1–3).

The Networks View

A second perspective on social capital, which attempts to account for both its upside and its downside, stresses the importance of vertical as well as horizontal associations between people and of relations within and among such organizational entities as community groups and firms. Building on work by Granovetter (1973), it recognizes that strong intracommunity ties give families and communities a sense of identity and common purpose (Astone and others 1999). This view also stresses, however, that without weak intercommunity ties, such as those that cross various social divides based on religion, class, ethnicity, gender, and socioeconomic status, strong horizontal ties can become a basis for the pursuit of narrow sectarian interests. In the recent popular literature, the former has been called "bonding" and the latter "bridging" social capital (Gittell and Vidal 1998). Different combinations of these dimensions, it is argued, are responsible for the range of outcomes that can be attributed to social capital. This more nuanced perspective, which we call the *networks* view, re-

gards the tension between social capital's virtues and vices as a defining property, one that explains in part why scholars and policymakers have been so persistently ambivalent about its potential as a theoretical construct and policy instrument.

The networks view of social capital is closely associated with Burt (1992, 1997, 1998); Fafchamps and Minten (1999); Massey (1998); Massey and Espinosa (1997); Portes (1995, 1997, 1998); and Portes and Sensenbrenner (1993). It is characterized by two key propositions. First, social capital is a double-edged sword. It can provide a range of valuable services for community members, ranging from baby-sitting and house-minding to job referrals and emergency cash. But there are also costs in that those same ties can place considerable noneconomic claims on members' sense of obligation and commitment, with negative economic consequences. Group loyalties may be so strong that they isolate members from information about employment opportunities, foster a climate of ridicule toward efforts to study and work hard, or siphon off hard-won assets (say, to support recent immigrants from the home country). Portes and Sensenbrenner (1993) cite the case of prosperous Asian immigrants who anglicized their names in order to divest themselves of communal obligations to subsequent cohorts. Second, the sources of social capital need to be distinguished from the consequences derived from them. Imputing only desirable outcomes to social capital, or equating them with it, ignores the possibility that these outcomes may be attained at another group's expense, that given outcomes may be suboptimal, or that desirable outcomes attained today come at the price of significant costs tomorrow.

These results have given rise to the logical conclusion that both strong intracommunity ties and weak extracommunity networks are needed to avoid making tautological claims regarding the efficacy of social capital. (Without this distinction, for example, it could be argued that successful groups are distinguished by their dense community ties, failing to consider the possibility that the same ties could be *preventing* success in another otherwise similar group.) Accordingly, the networks view argues that communities can be characterized by their endowments of these two dimensions of social capital and that different combinations of these dimensions account for the range of outcomes associated with social capital (table 1).

Furthermore, as community members' welfare changes over time, so too does the optimal calculus of costs and benefits associated with particular combinations of bonds and bridges. Poor entrepreneurs, for example, initially dependent on their

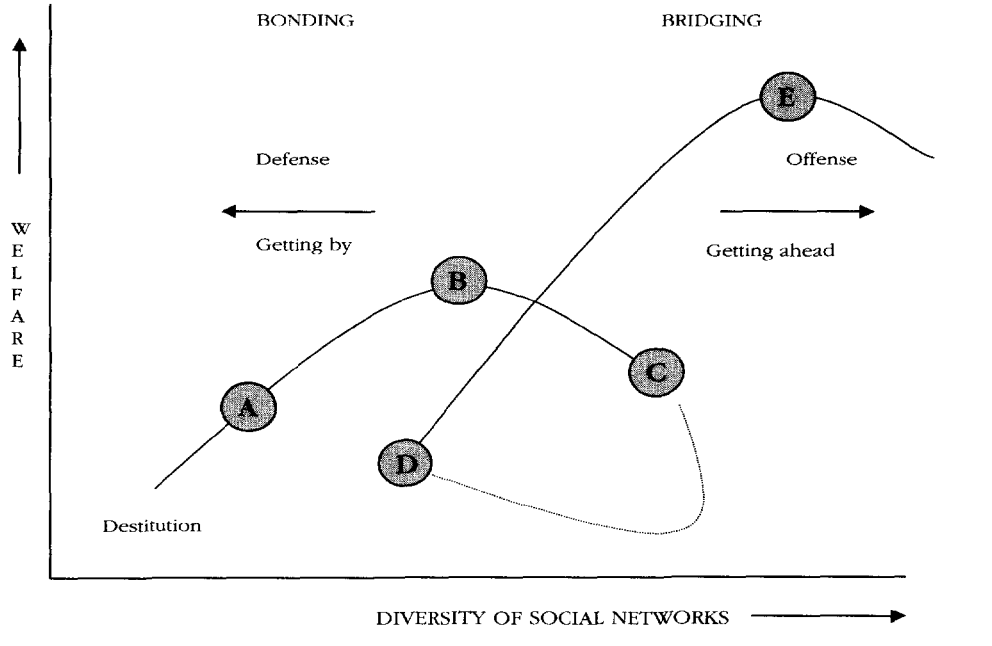
Table 1. *Dimensions of Social Capital at the Community Level*

Extracommunity networks (bridging)	*Intracommunity ties (bonding)*	
	Low	*High*
Low	Outcasts	Poor villagers
High	Recent rural-to-urban migrants	Successful members of microfinance programs

immediate neighbors and friends (their bonding social capital) for credit, insurance, and support, require access to more extensive product and factor markets as their businesses expand. Granovetter (1995) argues that economic development takes place through a mechanism that allows individuals to draw initially on the benefits of close community membership but that also enables them to acquire the skills and resources to participate in networks that transcend their community, thereby progressively joining the economic mainstream.

These insights can be demonstrated graphically and applied to poverty reduction more generally. Figure 1 shows that as the social networks of the poor become more diverse, so too does their welfare. The social capital residing in a given network can be leveraged or used more efficiently, which is essentially the genius of group-based credit programs such as the well-known Grameen Bank in Bangladesh (van Bastelaer 1999). Poor village women with no material collateral are given loans on the basis of their membership in a small peer group, which helps them start or expand a small business and thereby improve their families' welfare (A). But the economic returns

Figure 1. *Social Capital and Poverty Transitions*



Source: Woolcock (2000).

to any given group soon reach a limit (B), especially when they rely on high endowments of "bonding" social capital. If the group continues to expand—for example, through the arrival of subsequent cohorts from the village—its resources may become overwhelmed, thereby reducing the well-being of long-established members (C). Similarly, long-term members of group-based credit programs may find that obligations and commitments to their colleagues present obstacles to further advancement, especially for the more ambitious (Woolcock 1999). In these circumstances, many poor people partially divest themselves of their immediate community ties (D) and find a potentially more diverse network where "bridging" social capital is more abundant and economic opportunities more promising (E). Migration from villages to cities is the most dramatic example of this situation, but Portes and Sensenbrenner's (1993) name-changing Asian immigrants are doing essentially the same thing.

The networks view has been employed with great effect in recent development research. In their analysis of poor communities in rural areas of northern India, for example, Kozel and Parker (2000) report that social groups among poor villagers serve vitally important protection, risk management, and solidarity functions. It is the more extensive and leveraged networks of the nonpoor, by contrast, that are used for strategic advantage and the advancement of material interests. Crudely put, the networks of the poor play defense, while those of the nonpoor play offense. Barr (1998) reports strikingly similar results from her work on the relationship between the structure of business networks and enterprise performance in Africa. Poor entrepreneurs, operating small local firms in traditional industries, form what Barr calls solidarity networks to exchange personal information about members' conduct and intentions. The primary function of these networks is to reduce risk and uncertainty. Larger regional firms, in contrast, coalesce into innovation networks that share knowledge about technology and global markets with the explicit goal of enhancing enterprise profit, productivity, and market share (see also Van Dijk and Rabellotti 1997; Fafchamps and Minten 1999). Far from dismissing the vitality of traditional village groups in poor communities (the modernization view) or romanticizing it (the communitarian view), the networks view in effect recognizes that these groups can both help and hinder economic advancement.

The clear challenge to social capital theory, research, and policy from the networks perspective is thus to identify the conditions under which the many positive aspects of bonding social capital in poor communities can be harnessed and its integrity retained (and, if necessary, its negative aspects dissipated), while simultaneously helping the poor gain access to formal institutions and a more diverse stock of bridging social capital. This process is fraught with multiple dilemmas, however, especially for external nongovernmental organizations, extension services, and development agencies, because it may entail altering social systems that are the product of longstanding cultural traditions or of powerful vested interests.

The particular strength of the networks view is its willingness to engage in detailed policy discussions on the basis of compelling empirical evidence and detailed assessments of the veracity of competing explanations. This view, however, minimizes the "public good" nature of social groups, regarding any benefits of group activity as primarily the property of the particular individuals involved. Its proponents thus are highly skeptical of arguments that social capital can (or should) be measured across larger social aggregates, such as societies or nations (Portes 1998). Neither does the networks approach explicitly incorporate institutions at the societal level and their capacity to both shape and be shaped by local communities. To be sure, the networks perspective recognizes that weak laws and overt discrimination can undermine efforts by poor minorities to act in their collective interest, but the role communities play in shaping institutional performance generally, and the enormous potential of positive state-society relations in particular, are largely ignored.

The Institutional View

A third perspective of social capital, which we call the *institutional* view, argues that the vitality of community networks and civil society is largely the product of the political, legal, and institutional environment. Where the communitarian and networks perspectives largely treat social capital as an independent variable giving rise to various outcomes, both good and bad, the institutional view instead views social capital as a dependent variable. This approach argues that the very capacity of social groups to act in their collective interest depends on the quality of the formal institutions under which they reside (North 1990). It also stresses that the performance of states and firms themselves depends on their own internal coherence, credibility, and competence and on their external accountability to civil society.

Research from the institutional view has two variants, both of which have yielded remarkably complementary results. The first approach, described by Skocpol (1995, 1996), encompasses case studies based on comparative history and contends that it is wrong to argue that firms and communities thrive to the extent that governments retreat. On the contrary, Skocpol shows, civil society thrives to the extent that the state actively encourages it. Tendler's (1997) research on the political economy of decentralization in Brazil similarly stresses the importance of good government for making local programs work.

A second, and increasingly influential, approach relies on quantitative cross-national studies of the effects of government performance and social divisions on economic performance. This approach, pioneered by Knack and Keefer (1995, 1997), equates social capital with the quality of a society's political, legal, and economic institutions. Drawing on various indexes of institutional quality compiled by investment agencies and human rights groups, these studies show that items such as "generalized trust," "rule of law," "civil liberties," and "bureaucratic quality" are positively

associated with economic growth. In a recent review of this particular strand of the literature, Knack (1999:28) concludes that "social capital reduces poverty rates and improves, or at a minimum does not worsen, income inequality."

Collier and Gunning (1999) employ a variation of this view in their analyses of the causes of slow growth in Africa (see also Collier 1998, 1999; Temple 1998). Distinguishing between civic and government social capital, they show that slow growth occurs in societies with both high levels of ethnic fragmentation and weak political rights. Although Rodrik (1998, 1999) does not employ the terminology of social capital, he makes a similar argument, demonstrating that economies with divided societies and weak institutions for managing conflict respond sluggishly to shocks. Easterly (2000) also reports that societies able to generate and sustain a middle-class consensus are those most likely to produce stable and positive rates of growth. The related literature on social capabilities and development (Hall and Jones 1999; Temple and Johnson 1998) tells a similar story.

Several empirical and methodological questions can be raised about these studies, but in aggregate their message is loud and clear. Rampant corruption, frustrating bureaucratic delays, suppressed civil liberties, vast inequality, divisive ethnic tensions, and failure to safeguard property rights (to the extent that they exist at all) are major impediments to prosperity. In countries where these conditions prevail, there is little to show for well-intentioned efforts to build schools, hospitals, roads, and communications infrastructure or to encourage foreign investment (World Bank 1998). Investments in civic and government social capital are thus highly complementary to investments in more orthodox forms of capital accumulation.

The very strength of the institutional view in addressing macroeconomic policy concerns, however, is also a weakness in that it lacks a microeconomic component. Freedoms, rights, and liberties, for example, have to be secured by government. Coherent and competent bureaucracies may take decades to construct and may yield benefits more immediately suited to corporate interests than to those of the poor. In providing broad statistical evidence for the importance of social capital, the subtlety, richness, and enormous variation gleaned from case studies of individual countries and communities is lost, as are the voices of those most directly affected by weak public institutions: the poor.

The Synergy View

In recognition of this disconnect, a number of scholars have recently proposed what might be called a *synergy view*, which attempts to integrate the compelling work emerging from the networks and institutional camps. Although the synergy view traces its intellectual antecedents to earlier work in comparative political economy and anthropology, its most influential body of research was published in a special issue of *World Development* (1996). The contributors to this volume examined cases

from Brazil, India, Mexico, the Republic of Korea, and Russia in search of the conditions that foster developmental synergies—dynamic professional alliances and relationships between and within state bureaucracies and various actors in civil society.

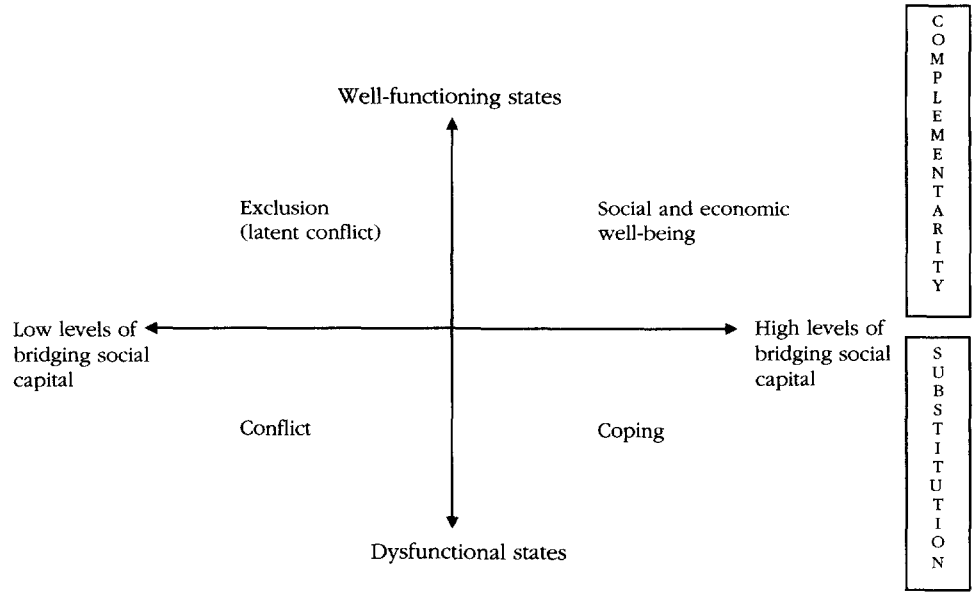
Three broad conclusions emerged from these studies:

- Neither the state nor societies are inherently good or bad; governments, corporations, and civic groups are variable in the impact they can have on the attainment of collective goals.
- States, firms, and communities alone do not possess the resources needed to promote broad-based, sustainable development; complementarities and partnerships forged both within and across these different sectors are required. Identifying the conditions under which these synergies emerge (or fail to emerge) is thus a central task of development research and practice.
- Of these different sectors, the state's role in facilitating positive developmental outcomes is the most important and problematic. This is so because the state is not only the ultimate provider of public goods (stable currencies, public health, universal education) and the final arbiter and enforcer of the rule of law (property rights, due process, freedom of speech and association) but is also the actor best able to facilitate enduring alliances across the boundaries of class, ethnicity, race, gender, politics, and religion. Communities and firms also have an important role to play in creating the conditions that produce, recognize, and reward good governance. In otherwise difficult institutional environments, community leaders who are able to identify and engage what Fox (1992) calls "pockets of efficiency within the state" become agents of more general reform.

Evans (1992, 1995, 1996), one of the primary contributors to this view, concludes that synergy between government and citizen action is based on *complementarity* and *embeddedness.* Complementarity refers to mutually supportive relations between public and private actors and is exemplified in legal frameworks that protect rights of association and in more humble measures such as chambers of commerce to facilitate exchanges among community associations and business groups. Embeddedness refers to the nature and extent of the ties connecting citizens and public officials. The classic examples are from irrigation, in which the lowest-level irrigation officials are from the community being served; they are enmeshed in local social relations and hence are under pressure by the community to perform and be responsive to them. Importantly, this approach works only where the actions of public officials are simultaneously bound by performance-oriented organizational environments that are competent, coherent, and credible. As the case of Russia amply demonstrates, weak public institutions and deep cleavages between powerful authorities and ordinary citizens can lead to political instability, rampant corruption, rising inequality, and capital flight (Rose 1998).

Developing these ideas, Woolcock (1998) shows that a range of development outcomes flows from different types and combinations of community capacity and state functioning. Narayan (1999) integrates the core ideas of bridging social capital and state-society relations and suggests that different interventions are needed for different combinations of governance and bridging social capital in a group, community, or society (figure 2). In societies (or communities) with good governance and high levels of bridging social capital, there is complementarity between state and society, and economic prosperity and social order are likely. But when a society's social capital inheres mainly in primary social groups disconnected from one another, the more powerful groups dominate the state, to the exclusion of other groups. Such societies, which include countries in Latin America with large excluded indigenous populations, are characterized by latent conflict. In these circumstances, a key task for subordinate groups and activists is to forge broad, coherent coalitions (Keck and Sikkink 1998) and nurture relations with allies in positions of power (Fox and Brown 1998); should they be successful, weak groups may begin to accrue rights and resources previously denied them. Similarly, a state that opens up and explicitly builds bridges

Figure 2. *Relationship between Bridging Social Capital and Governance*



Note: Complementarity refers to the optimal interaction of government and markets in civil society; substitution is the replacement by informal organizations (families, networks, and so on) of services ordinarily provided by governments and institutions.

Source: Adapted from Narayan (1999).

to excluded groups increases the likelihood that the poor will be able to gain access to the resources and services to which they are entitled.

Alternatively, state-society relations may degenerate into conflict, violence, war or anarchy—a breakdown that allows warlords, local mafias, and guerrilla movements to take over the power and authority of the state. Restoring economic prosperity and peace to Rwanda, for example, will involve forging a measure of reconciliation between two ethnic groups. Often, when citizens are deprived of services and benefits, informal networks substitute for the failed state and form the basis of coping strategies. This is the case in Benin and Togo, where women, denied access to formal credit, established informal revolving credit societies; in Tanzania the absence of police protection has led some villages to rely on their own system of security guards (Narayan and others 2000).

When representatives of the state, the corporate sector, and civil society establish common forums through which they can pursue common goals, development can proceed. In these circumstances social capital has a role as a mediating variable that is shaped by public and private institutions. This shaping is an inherently contentious and political process, one in which the role of the state is crucial. Moreover, the fundamental social transformation of economic development—from traditional kinship-based community life to societies organized by formal institutions—alters the calculus of costs and benefits associated with different dimensions of social capital and the desirable combinations of these dimensions (Berry 1993). Although development struggles are inherently political, they are not always won by the most powerful, nor do challenges to authority always entail violent conflict. Patient efforts by intermediaries to establish partnerships between associations of the poor and outsiders can reap significant dividends (Isham, Narayan, and Pritchett 1995). As Uphoff (1992:273) points out,

> paradoxical though it may seem, "top-down" efforts are usually needed to introduce, sustain, and institutionalize "bottom-up" development. We are commonly constrained to think in "either-or" terms—the more of one the less of the other—when both are needed in a positive-sum way to achieve our purposes.

The synergy view suggests three central tasks for theorists, researchers, and policymakers: to identify the nature and extent of a community's social relationships and formal institutions, and the interaction between them; to develop institutional strategies based on these social relations, particularly the extent of bonding and bridging social capital; and to determine how the positive manifestations of social capital—cooperation, trust, and institutional efficiency—can offset sectarianism, isolationism, and corruption. Put another way, the challenge is to transform situations where a community's social capital substitutes for weak, hostile, or indifferent formal institutions into ones in which both realms complement one another.

Table 2. *Four Views of Social Capital*

Perspective	*Actors*	*Policy prescriptions*
Communitarian view		
Local associations	Community groups	Small is beautiful
	Voluntary organizations	Recognize social assets of the poor
Networks view		
Bonding and bridging community ties	Entrepreneurs	Decentralize
	Business groups	Create enterprise zones
	Information brokers	Bridge social divides
Institutional view		
Political and legal institutions	Private and public sectors	Grant civil and political liberties
		Institute transparency, accountability
Synergy view		
Community networks and state-society relations	Community groups, civil society, firms, states	Coproduction, complementarity
		Participation, linkages
		Enhance capacity and scale of local organizations

Table 2 summarizes the key elements of the four perspectives on social capital and development and their corresponding policy prescriptions. The differences between them are primarily the unit of analysis on which they focus; their treatment of social capital as an independent, dependent, or mediating variable; and the extent to which they incorporate a theory of the state. The largest and most influential bodies of work have emerged from the networks and institutional perspectives; the most recent approaches seek a synthesis.

Measuring Social Capital

Several recent innovative studies have attempted to quantify social capital and its contribution to economic development. To arrive at concrete policy recommendations for using social capital as a development tool, more comparative research is required that uses precise measures of social capital to examine within-country and across-country variations in poverty reduction, government performance, ethnic conflict, and economic growth. Obtaining a single, true measure of social capital is probably not possible, for several reasons. First, the most comprehensive definitions of social capital are multidimensional, incorporating different levels and units of analysis. Second, the nature and forms of social capital change over time, as the balance shifts between informal organizations and formal institutions. And third, because no long-standing cross-country surveys were initially designed to measure social capital, contemporary researchers have had to compile indexes from a range of approximate items (measures of trust, confidence in government, voting trends, social mobility,

and so on). Several excellent studies have identified useful measures of, and proxies for, social capital, however.

One measure is membership in informal and formal associations and networks. In developing countries generally, and in rural areas in particular, measures that capture the informal give-and-take through communitywide festivals, sporting events, and other traditional methods of fostering social connections are very important indicators of the underlying stocks of social capital. Based on data from a survey of 1,400 households in 87 villages across Tanzania (Narayan 1997), Narayan and Pritchett (1999) developed an index of social capital at the household and community levels that included density and characteristics of informal and formal groups and networks. The dimensions of this index included group functioning, financial and in-kind contributions to groups, participation in decisionmaking, and heterogeneity of membership. A series of measures was also constructed on interpersonal trust and changes over time. These measures demonstrated that social capital was indeed both social and capital, generating returns that exceeded those to human capital.

In tandem with the Tanzania study, studies of local institutions in three countries—Bolivia (Grootaert and Narayan 2000), Burkina Faso (Grootaert, Oh, and Swamy 1999), and Indonesia (Grootaert 1999)—looked at qualitative service delivery issues and quantified these variables. These studies demonstrated that the questionnaire items do in fact capture different dimensions of social capital at the household and community levels, that certain dimensions of social capital contribute significantly to household welfare, and that social capital is the capital of the poor. The most important variables in these studies are density of associations, heterogeneity of membership in associations, and degrees of active participation in them.

Another manifestation of social capital includes norms and values that facilitate exchanges, lower transaction costs, reduce the cost of information, permit trade in the absence of contracts, and encourage responsible citizenship and the collective management of resources (Fukuyama 1995). Inglehart's (1997) work on the World Values Survey is the most comprehensive effort in this area. The questions economists working on social capital find valuable are those on trust ("Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?"). Knack and Keefer (1997), for example, use these data to show the positive relationship between trust and levels of investment in a country.

Although research attempting to identify the nature of the relationships between social variables and development has recently proliferated, the quality of the data is less than ideal. With mounting pressure to provide simple measures of inherently complex and interdependent relationships, there is a danger that expectations will exceed capacity and that hastily assembled, poorly conceived measures will jeopardize the agenda they purport to serve. One way to strike the balance between quality and quantity measures is to unbundle social capital into its dimensions and to generate new data sets that are comparable across many countries.[2]

Four recent studies attempt to develop indices of social capital at the national or subnational levels. In the United States several new surveys of civic engagement are being conducted in addition to the work already collected in surveys of consumer preferences and changes in lifestyles. The National Commission on Philanthropy and Civic Renewal (1998), for example, has developed a National Index of Civic Engagement based on a sample of 1,000 respondents. This index includes five dimensions: the giving climate, community engagement, charitable involvement, the spirit of voluntarism, and active citizenship. Robert Putnam's Saguaro Seminar will soon launch the Social Capital Community Benchmark, a comprehensive survey of social capital in the United States (Putnam 2000).

In exploring the roots and determinants of Hindu and Muslim riots in India, Varshney (2000) focuses on the role of intercommunal networks. In cities where Hindus and Muslims have little interaction, Varshney shows that latent communal conflict has few channels for peaceful resolution and periodically descends into violence; in cities where association memberships overlap and everyday interactions are frequent, conflict is anticipated and dissipated. This research was based on six Indian cities carefully arranged in three matched pairs that were similar in terms of Hindu-Muslim demographic composition but dissimilar in that one city experienced recurrent riots whereas the other city remained calm. Varshney's work shows that diversity can be a source of strength where social ties transcend different community boundaries.

To assess social capital at the community level, Onyx and Bullen (forthcoming) developed a questionnaire for the state of New South Wales, Australia, from which they isolated eight underlying factors that constituted an individual's social capital: participation in the local community, proaction in a social context, feelings of trust and safety, neighborhood connections, connections with family and friends, tolerance of diversity, value of life, and work connections. Looking only at an individual's social capital score, the authors could predict the community to which the person belonged, thus raising the prospects for this instrument being used for planning and monitoring community development activities.

Building on this work, researchers are working to develop social capital instruments that can be used as diagnostic tools at the community level and across countries. Because the forms of social capital are society-specific and change over time, the instruments must focus on a range of dimensions of social capital (Narayan and Cassidy 1999). Such instruments have recently been introduced in Ghana and Uganda (Narayan 1998) and by the World Bank's Social Capital Initiative in Panama and India (Krishna and Shrader 1999).[3] Analyses of the data reveal that the dimensions underlying social capital are strikingly similar even when the context is quite different. The Ghana study draws on a sample of 1,471 rural and urban households, while the Uganda study focuses on 950 households in slum communities in Kampala. Factor analyses reveal a similar underlying structure and clustering of variables.

Implications for Development Theory and Policy

The concept of social capital offers a way to bridge sociological and economic perspectives and to provide potentially richer and better explanations of economic development. One important way it does this is by showing that the nature and extent of social interactions between communities and institutions shape economic performance. This, in turn, has important implications for development policy, which has long focused exclusively on an economic dimension. Similarly, understanding how outside agencies can work to alleviate poverty in diverse and poorly understood communities remains one of the great challenges of development. A social capital perspective stresses that technical and financial soundness is a necessary but insufficient condition for acceptance of a project by poor communities.

Six broad recommendations can be offered for incorporating the concept of social capital into development policy. First, for development interventions in all sectors and at all levels (especially the country level), social institutional analysis should be used to identify correctly the range of stakeholders and their interrelations. Understanding how proposed policy interventions will affect the power and political interests of the stakeholders is a vital consideration, since all policy interventions occur in a social context characterized by a delicate mix of informal organizations, networks, and institutions. The design of an intervention needs to pay special attention to the potential for dominant groups to mobilize in ways that undermine the public good.

Second, it is critical to invest in the organizational capacity of the poor and to help build bridges between communities and social groups. The latter is particularly important because many decisions affecting the poor are not made at the local level. To this end, the use of participatory processes can facilitate consensus-building and social interaction among stakeholders with diverse interests and resources. Finding ways and means by which to transcend social divides and build social cohesion and trust is crucial for economic development. One of the great virtues of the idea and discourse on social capital is that it provides a common language for these different stakeholders, enabling them to communicate more easily with one another.

Third, a social capital perspective adds its voice to those calling for information disclosure policies at all levels to encourage informed citizenship and accountability of both private and public actors who purport to serve the public good. Fourth, improvements in physical access and modern communications technology that can foster information exchange across social groups should be emphasized to complement social interaction based on face-to-face interchange. Fifth, development interventions should be viewed through a social capital lens, and assessments of their impact should include the potential effects of the intervention on the social capital of poor communities. To reiterate, the social networks of the poor are one of the primary resources they have for managing risk and vulnerability, and outside agents therefore need to find ways to complement these resources, rather than substitute for them.

Finally, social capital should be seen as a component of orthodox development projects, from dams and irrigation systems to local schools and health clinics. Where poor communities have direct input into the design, implementation, management, and evaluation of projects, returns on investments and the sustainability of the project are enhanced (Esman and Uphoff 1984).

Conclusion

Although it is too soon to announce the arrival of a new development paradigm, it is not unreasonable to claim that a consensus is emerging about the importance of social relations in development. In unpacking the literature on social capital and development, a recurring message is that social relations provide opportunities for mobilizing other growth-enhancing resources, that social capital does not exist in a political vacuum, and that the nature and extent of the interactions between communities and institutions hold the key to understanding the prospects for development in a given society. The evidence supports the argument that social capital can be used to promote or to undermine the public good. This consideration suggests that one of the most important examples of social capital at work in the absence of formal insurance mechanisms and financial instruments is the use by the poor of social connections to protect themselves against risk and vulnerability.

In many respects the research on social capital is still in its early stages, but practitioners and policymakers cannot wait for researchers to know all there is to know before acting. Instead, all those involved should adopt a stance of learning by doing. This implies more rigorous evaluations of project and policy impact on social capital, more work on unbundling the mechanisms through which social capital works, and understanding the determinants of social capital itself. It also implies that practical lessons emerging from development projects can themselves be used to inform social capital theory.

It would be the ultimate irony if those people most interested in studying social capital and promoting its use in formulating development policy did not themselves foster trust, openness, and a willingness to share information, ideas, and opportunities in this field. Readers are invited to access, use, and contribute to the ongoing research on social capital.[4] It is only through collaborative efforts—with all that this entails regarding struggle, perseverance, negotiation, and mutual willingness to learn—that genuine progress will be made.

Notes

Michael Woolcock is a social scientist with the World Bank's Development Research Group and an adjunct lecturer in public policy at Harvard University. Deepa Narayan is a lead social development

specialist in the Poverty Reduction and Economic Management Network at the World Bank. For helpful comments on an earlier draft of this paper, the authors thank John Blaxall, Jonathan Fox, Christiaan Grootaert, Bill Mulford, Vijayendra Rao, Anders Rudkvist, and anonymous reviewers.

1. For citations on the first eight fields, see Woolcock (1998) and Foley and Edwards (1999). See also the database of articles on the World Bank's social capital Website, at <http://www.worldbank.org/poverty/scapital/library/index.htm#db>.

2. A number of recent survey instruments are available to researchers doing work in this field. See the World Bank's social capital Website, <http://www.worldbank.org/poverty/scapital/library/surveys.htm>.

3. The World Bank's Social Capital Initiative is a $1.2 million dollar project sponsored by the government of Denmark. Several monographs produced for the initiative have been cited in this paper; these and several others can be downloaded at <http://www.worldbank.org/poverty/scapital/wkrppr/wrkppr.htm>. These papers are currently being edited and prepared for formal publication.

4. The World Bank's Social Capital Thematic Group Website contains instructions on how to receive our newsletter and join the e-mail discussion group. Go to <http://www.worldbank.org/poverty/scapital/>.

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Astone, Nan Marie, Constance Nathanson, Robert Schoen, and Young Kim. 1999. "Family Demography, Social Theory, and Investment in Social Capital." *Population and Development Review* 25(1):1–31.

Barr, Abigail. 1998. "Enterprise Performance and the Functional Diversity of Social Capital." Working Paper Series 98-1. University of Oxford, Institute of Economics and Statistics, Oxford, U.K. Processed.

Berry, Sara. 1993. *No Condition Is Permanent: The Social Dynamics of Agrarian Change in Sub-Saharan Africa.* Madison, Wis.: University of Wisconsin Press.

Briggs, Xavier de Souza. 1998. "Brown Kids in White Suburbs: Housing Mobility and the Multiple Faces of Social Capital." *Housing Policy Debate* 9(1):177–221.

Brown, L. David. 1998. "Creating Social Capital: Nongovernmental Development Organizations and Intersectoral Problem Solving." In Walter W. Powell and Elisabeth Clemens, eds., *Private Action and the Public Good.* New Haven, Conn.: Yale University Press.

Brown, L. David, and Darcy Ashman. 1996. "Participation, Social Capital, and Intersectoral Problem Solving: African and Asian Cases." *World Development* 24(6):1477–95.

Burt, Ronald. 1992. *Structural Holes: The Social Structure of Competition.* Cambridge, Mass.: Harvard University Press.

———. 1997. "The Contingent Value of Social Capital." *Administrative Science Quarterly* 42:339–65.

———. 1998. "The Gender of Social Capital." *Rationality and Society* 10(1):5-46.

Coleman, James. 1987. "Norms as Social Capital." In Gerard Radnitzky and Peter Bernholz, eds., *Economic Imperialism: The Economic Method Applied outside the Field of Economics.* New York: Paragon House Publishers.

———. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94(Supplement):S95–S120.

———. 1990. *Foundations of Social Theory.* Cambridge, Mass.: Harvard University Press.

Collier, Paul. 1998. "Social Capital and Poverty." Social Capital Initiative Working Paper 4. World Bank, Social Development Department, Washington, D.C. Processed.

———. 1999. "The Political Economy of Ethnicity." In Boris Pleskovic and Joseph E. Stiglitz, eds., *Annual World Bank Conference on Development Economics 1998.* Washington, D.C.: World Bank.

Collier, Paul, and Jan Willem Gunning. 1999. "Explaining African Economic Performance." *Journal of Economic Literature* 37(March):64–111.

Dordick, Gwendolyn. 1997. *Something Left to Lose: Personal Relations and Survival among New York's Homeless.* Philadelphia, Pa.: Temple University Press.

Easterly, William. 2000. "Happy Societies: The Middle Class Consensus and Economic Development." World Bank, Development Research Group, Washington, D.C. Processed.

Escobar, Arturo. 1995. *Encountering Development: The Making and Unmaking of the Third World.* Princeton, N.J.: Princeton University Press.

Esman, Milton, and Norman Uphoff. 1984. *Local Organizations: Intermediaries in Rural Development.* Ithaca, N.Y.: Cornell University Press.

Evans, Peter. 1992. "The State as Problem and Solution: Predation, Embedded Autonomy, and Structural Change." In Stephan Haggard and Robert Kaufman, eds., *The Politics of Economic Adjustment: International Constraints, Distributive Conflicts, and the State.* Princeton, N.J.: Princeton University Press.

———. 1995. *Embedded Autonomy: States and Industrial Transformation.* Princeton, N.J.: Princeton University Press.

———. 1996. "Government Action, Social Capital and Development: Reviewing the Evidence on Synergy." *World Development* 24(6):1119–32.

Fafchamps, Marcel, and Bart Minten. 1999. "Social Capital and the Firm: Evidence from Agricultural Trade." Social Capital Initiative Working Paper 17. World Bank, Social Development Department, Washington, D.C. Processed.

Foley, Michael, and Bob Edwards. 1999. "Is It Time to Disinvest in Social Capital?" *Journal of Public Policy* 19:141–73.

Fox, Jonathan. 1992. "Democratic Rural Development: Leadership Accountability in Regional Peasant Organizations." *Development and Change* 23(2):1–36.

Fox, Jonathan, and L. David Brown, eds. 1998. *The Struggle for Accountability: The World Bank, NGOs, and Grassroots Movements.* Cambridge, Mass.: MIT Press.

Fukuyama, Francis. 1995. *Trust: The Social Virtues and the Creation of Prosperity.* New York: Free Press.

Gittell, Ross, and Avis Vidal. 1998. *Community Organizing: Building Social Capital as a Development Strategy.* Newbury Park, Calif.: Sage Publications.

Granovetter, Mark. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360–80.

———. 1995. "The Economic Sociology of Firms and Entrepreneurs." In Alejandro Portes, ed., *The Economic Sociology of Immigration: Essays on Networks, Ethnicity, and Entrepreneurship.* New York: Russell Sage Foundation.

Grootaert, Christiaan. 1999. "Social Capital, Household Welfare, and Poverty in Indonesia." Local Level Institutions Working Paper 6. World Bank, Social Development Department, Washington, D.C. Processed.

Grootaert, Christiaan, and Deepa Narayan. 2000. "Local Institutions, Poverty, and Household Welfare in Bolivia." Working Paper 9. World Bank, Social Development Department, Washington, D.C. Processed.

Grootaert, Christiaan, Gi-Taik Oh, and Anand Swamy. 1999. "Social Capital and Development Outcomes in Burkina Faso." Working Paper 7. World Bank, Social Development Department, Washington, D.C. Processed.

Hall, Robert, and Charles Jones. 1999. "Why Do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114(1):83–116.

Hanifan, Lyda J. 1916. "The Rural School Community Center." *Annals of the American Academy of Political and Social Science* 67:130–38.

Holzmann, Robert, and Steen Jorgensen. 1999. "Social Protection as Social Risk Management." Social Protection Discussion Paper 9904. World Bank, Human Development Network, Social Protection Team, Washington, D.C. Processed.

Homans, George. 1961. *Social Behavior: Its Elementary Forms.* New York: Harcourt, Brace and World.

Inglehart, Ronald. 1997. *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies.* Princeton, N.J.: Princeton University Press.

Isham, Jonathan. 1999. "The Effect of Social Capital on Technology Adoption: Evidence from Rural Tanzania." Paper presented at the annual meeting of the American Economic Association, New York. Processed.

Isham, Jonathan, and Daniel Kaufmann. 1999. "The Forgotten Rationale for Policy Reform: The Productivity of Investment Projects." *Quarterly Journal of Economics* 114(1):149–84.

Isham, Jonathan, Deepa Narayan, and Lant Pritchett. 1995. "Does Participation Improve Performance? Establishing Causality with Subjective Data." *The World Bank Economic Review* 9(2):175–200.

Jacobs, Jane. 1961. *The Life and Death of Great American Cities.* New York: Random House.

Keck, Margaret, and Kathryn Sikkink. 1998. *Activists beyond Borders: Advocacy Networks in International Politics.* Ithaca, N.Y.: Cornell University Press.

Knack, Stephen. 1999. "Social Capital, Growth, and Poverty: A Survey of Cross-Country Evidence." Social Capital Initiative Working Paper 7. World Bank, Social Development Department, Washington, D.C. Processed.

Knack, Stephen, and Philip Keefer. 1995. "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures." *Economics and Politics* 7(3):207–27.

———. 1997. "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation." *Quarterly Journal of Economics* 112:1251–88.

Kozel, Valerie, and Barbara Parker. 2000. "Integrated Approaches to Poverty Assessment in India." In Michael Bamberger, ed., *Integrating Quantitative and Qualitative Research in Development Projects.* Washington, D.C.: World Bank.

Krishna, Anirudh, and Elizabeth Shrader. 1999. "Social Capital Assessment Tool." Social Capital Initiative Working Paper 22. World Bank, Social Development Department, Washington, D.C. Processed.

Loury, Glenn C. 1977. "A Dynamic Theory of Racial Income Differences." In P. A. Wallace and A. LeMund, eds., *Women, Minorities, and Employment Discrimination.* Lexington, Mass.: Lexington Books.

Massey, Douglas. 1998. "March of Folly: U.S. Immigration Policy after NAFTA." *The American Prospect* 37:22–33.

Massey, Douglas, and Karin Espinosa. 1997. "What's Driving Mexico-U.S. Migration? A Theoretical, Empirical, and Policy Analysis." *American Journal of Sociology* 102(4):939–99.

Moore, Mick. 1997. "Societies, Polities, and Capitalists in Developing Countries: A Literature Survey." *Journal of Development Studies* 33(3):287–363.

Moser, Caroline. 1996. *Confronting Crisis: A Comparative Study of Household Responses to Poverty and Vulnerability in Four Poor Urban Communities.* Environmentally Sustainable Development Studies and Monographs Series 8. Washington, D.C.: World Bank.

Narayan, Deepa. 1995. "Designing Community-Based Development." Social Development Paper 7. World Bank, Environmentally and Socially Sustainable Development Network, Washington, D.C. Processed.

———. 1997. "Voices of the Poor: Poverty and Social Capital in Tanzania." Environmentally Sustainable Development Monograph 20. World Bank, Washington, D.C.

———. 1998. "Social Capital Survey in Ghana—Preliminary Results." World Bank, Poverty Reduction and Economic Management Network, Washington, D.C. Processed.

———. 1999. "Bonds and Bridges: Social Capital and Poverty." Policy Research Working Paper 2167. World Bank, Poverty Reduction and Economic Management Network, Washington, D.C. Processed.

Narayan, Deepa, and Michael Cassidy. 1999. "A Dimensional Approach to Measuring Social Capital: Development and Validation of a Social Capital Inventory." World Bank, Poverty Reduction and Economic Management Network, Washington, D.C. Processed.

Narayan, Deepa, and David Nyamwaya. 1996. "Learning from the Poor: A Participatory Poverty Assessment in Kenya." Environment Department Papers, Participation Series 34. World Bank, Social Policy and Resettlement Division, Washington, D.C. Processed.

Narayan, Deepa, and Lant Pritchett. 1999. "Cents and Sociability: Household Income and Social Capital in Rural Tanzania." *Economic Development and Cultural Change* 47(4):871–97.

Narayan, Deepa, and Talat Shah. 1999. "Gender Inequity, Poverty, and Social Capital." Background paper prepared for the Policy Research Report on Gender and Development. World Bank, Development Research Group, Washington, D.C. Processed.

Narayan, Deepa, with Raj Patel, Kai Schafft, Anne Rademacher, and Sarah Koch-Schulte. 2000. *Voices of the Poor. Can Anyone Hear Us? Voices from 46 Countries.* Washington, D.C.: World Bank.

National Commission on Philanthropy and Civic Renewal. 1998. "National Index of Civic Engagement." University of Connecticut, Storrs, Conn.

North, Douglass C. 1990. *Institutions, Institutional Change, and Economic Performance.* New York: Cambridge University Press.

Onyx, Jenny, and Paul Bullen. Forthcoming. "Measuring Social Capital in Five Communities." *Journal of Applied Behavioral Science.*

Platteau, Jean-Philippe. 1994. "Behind the Market Stage Where Real Societies Exist" (Parts I and II). *Journal of Development Studies* 30:533–77 and 753–817.

Portes, Alejandro, ed. 1995. *The Economic Sociology of Immigration: Essays on Networks, Ethnicity, and Entrepreneurship.* New York: Russell Sage Foundation.

———. 1997. "Neoliberalism and the Sociology of Development." *Population and Development Review* 23(2):229–59.

———. 1998. "Social Capital: Its Origins and Applications in Contemporary Sociology." *Annual Review of Sociology* 24:1–24.

Portes, Alejandro, and Patricia Landolt. 1996. "The Downside of Social Capital." *The American Prospect* 26 (May–June):18–21, 94.

Portes, Alejandro, and Julia Sensenbrenner. 1993. "Embeddedness and Immigration: Notes on the Social Determinants of Economic Action." *American Journal of Sociology* 98(6):1320–50.

Putnam, Robert. 1993. *Making Democracy Work: Civic Traditions in Modern Italy.* Princeton, N.J.: Princeton University Press.

———. 1995. "Tuning In, Tuning Out: The Strange Disappearance of Social Capital in America." *PS: Political Science and Politics* (December):664–83.

———. 2000. *Bowling Alone: The Collapse and Revival of American Community.* New York: Simon and Schuster.

Rodrik, Dani. 1998. "Where Did All the Growth Go? External Shocks, Social Conflict, and Growth Collapses." NBER Working Paper 6350. National Bureau of Economic Research, Cambridge, Mass.

———. 1999. *Making Openness Work.* Baltimore, Md.: Johns Hopkins University Press.

Rose, Richard. 1998. "Getting Things Done in an Anti-Modern Society: Social Capital Networks in Russia." Social Capital Initiative Working Paper 8. World Bank, Social Development Department, Washington, D.C. Processed.

Rubio, Mauricio. 1997. "Perverse Social Capital: Some Evidence from Colombia." *Journal of Economic Issues* 31(3):805–16.

Schafft, Kai. 1998. "Grassroots Development and the Reconfiguration of Local Political Institutions: Local Minority Self-Governance as a Political and Economic Resource for Hungary's Roma Population." Paper presented at the Graduate Student Conference on International Affairs, George Washington University, Washington, D.C. Processed.

Seeley, John R., Alexander R. Sim, and Elizabeth W. Loosley. 1956. *Crestwood Heights: A Study of the Culture of Suburban Life.* New York: Basic Books.

Skocpol, Theda. 1995. *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States.* Cambridge, Mass.: Harvard University Press.

———. 1996. "Unraveling from Above." *The American Prospect* 25:20–25.

Temple, Jonathan. 1998. "Initial Conditions, Social Capital, and Growth in Africa." *Journal of African Economies* 7(3):309–47.

Temple, Jonathan, and Paul Johnson. 1998. "Social Capability and Economic Growth." *Quarterly Journal of Economics* 113(3):965–90.

Tendler, Judith. 1997. *Good Government in the Tropics.* Baltimore, Md.: Johns Hopkins University Press.

United Nations. 1951. *Measures for the Economic Development of Underdeveloped Countries.* New York: United Nations, Department of Social and Economic Affairs.

Uphoff, Norman. 1992. *Learning from Gal Oya: Possibilities for Participatory Development and Post-Newtonian Social Science.* Ithaca, N.Y.: Cornell University Press.

van Bastalaer, Thierry. 1999. "Does Social Capital Facilitate the Poor's Access to Credit? A Review of the Microeconomic Literature." Social Capital Initiative Working Paper 8. World Bank, Social Development Department, Washington, D.C. Processed.

Van Dijk, Meine Pieter, and Roberta Rabellotti. 1997. *Enterprise Clusters and Networks in Developing Countries.* London: Frank Cass.

Varshney, Ashutosh. 2000. *Ethnic Conflict and Civic Life: Hindus and Muslims in India.* New Haven, Conn.: Yale University Press.

White, T. Anderson, and Glenn Smucker. 1998. "Social Capital and Governance in Haiti: Traditions and Trends." In *The Challenges of Poverty Reduction.* Report 17242-HA. Washington, D.C.: World Bank.

Wilson, William Julius. 1987. *The Truly Disadvantaged.* Chicago, Ill.: University of Chicago Press.

———. 1996. *When Work Disappears: The World of the New Urban Poor.* New York: Knopf.

Woolcock, Michael. 1998. "Social Capital and Economic Development: Toward a Theoretical Synthesis and Policy Framework." *Theory and Society* 27(2):151–208.

———. 1999. "Learning from Failures in Microfinance: What Unsuccessful Cases Tell Us about How Group-Based Programs Work." *American Journal of Economics and Sociology* 58(1):17–42.

———. 2000. "Managing Risk, Shocks, and Opportunities in Developing Economies: The Role of Social Capital." In Gustav Ranis, ed., *Dimensions of Development.* New Haven, Conn.: Yale Center for International and Area Studies.

———. Forthcoming. *Using Social Capital: Getting the Social Relations Right in the Theory and Practice of Economic Development.* Princeton, N.J.: Princeton University Press.

World Bank. 1989. "Staff Appraisal Report, Rwanda, Agricultural Services Project." Report 7599-RW. Agricultural Operations Division, South-Central and Indian Ocean Department, Africa Region. Processed.

———. 1998. *Assessing Aid: What Works, What Doesn't, and Why.* New York: Oxford University Press.

World Development. 1996. Special issue on social capital. 24(6).

What Should the World Bank Think about the Washington Consensus?

John Williamson

The phrase "Washington Consensus" has become a familiar term in development policy circles in recent years, but it is now used in several different senses, causing a great deal of confusion. In this article the author distinguishes between his original meaning as a summary of the lowest common denominator of policy advice addressed by the Washington-based institutions (including the World Bank) and subsequent use of the term to signify neoliberal or market-fundamentalist policies. He argues that the latter policies could not be expected to provide an effective framework for combating poverty but that the original advice is still broadly valid. The article discusses alternative ways of addressing the confusion. It argues that any policy manifesto designed to eliminate poverty needs to go beyond the original version but concludes by cautioning that no consensus on a wider agenda currently exists.

Ten years ago I invented the term "Washington Consensus" to refer to the lowest common denominator of policy advice being addressed by the Washington-based institutions to Latin American countries as of 1989 (Williamson 1990). While it is jolly to become famous for coining a term that reverberates around the world, I have long been doubtful about whether my phrase served to advance the cause of rational economic policymaking. My initial concern was that the phrase invited the interpretation that the liberalizing economic reforms of the past two decades were imposed by Washington-based institutions (for example, see Stewart 1997) rather than having resulted from the process of intellectual convergence that I believe underlies the reforms.[1] Richard Feinberg's "universal convergence" (in Williamson 1990) or Jean Waelbroeck's "one-world consensus" (Waelbroeck 1998) would have been a much better term for the intellectual convergence that I had in mind.

I have gradually developed a second and more significant concern, however. I find that the term has been invested with a meaning that is significantly different from that which I had intended and is now used as a synonym for what is often called

The World Bank Research Observer, vol. 15, no. 2 (August 2000), pp. 251–64.

"neoliberalism" in Latin America, or what Geeorge Soros (1998) has called "market fundamentalism." When I first came across this usage, I asserted that it was a misuse of my intended meaning. I had naïvely imagined that just because I had invented the expression, I had some sort of intellectual property rights that entitled me to dictate its meaning, but in fact the concept had become public property.

The battle of economic ideas, as McCloskey (1998) has argued, is fought to a significant extent with rhetoric. The use of a term with dual meanings and strong ideological overtones can therefore pose serious dangers not only of misunderstanding but also of inadvertently prejudicing policy objectives. Specifically, there is a real danger that many of the economic reforms favored by international development institutions—notably macroeconomic discipline, trade openness, and market-friendly microeconomic policies—will be discredited in the eyes of many observers, simply because these institutions are inevitably implicated in views that command a consensus in Washington and the term "Washington Consensus" has come to be used to describe an extreme and dogmatic commitment to the belief that markets can handle everything.

The objective of this article is to consider what should be done to minimize the damage to the cause of intellectual understanding, and therefore of rational economic reform, that is being wrought by the current widespread use of the term "Washington Consensus" in a sense different from that originally intended. Would it be productive, for example, to insist that the original usage is the correct one? Or should one simply refuse to debate in these terms? Is it possible to escape by declaring fidelity to some "post–Washington Consensus"? The first stage in answering these questions is a careful examination of the semantic issues involved.

The Original Version

My original paper (Williamson 1990) argued that the set of policy reforms that most of official Washington thought would be good for Latin American countries could be summarized in 10 propositions:

- Fiscal discipline
- A redirection of public expenditure priorities toward fields offering both high economic returns and the potential to improve income distribution, such as primary health care, primary education, and infrastructure
- Tax reform (to lower marginal rates and broaden the tax base)
- Interest rate liberalization
- A competitive exchange rate
- Trade liberalization
- Liberalization of inflows of foreign direct investment

- Privatization
- Deregulation (to abolish barriers to entry and exit)
- Secure property rights.

The need for the first three reforms is, so far as I am aware, widely accepted among economists. Nevertheless, when I reviewed the progress that Latin American countries had made in implementing the recommended set of policies several years later (Williamson 1996), it appeared that the least progress had come in redirecting public expenditure priorities. The other seven reforms have stimulated a measure of controversy and therefore merit comment.

In my original paper I specified interest rate liberalization as the fourth reform. I am now well aware that many economists have reservations about that formulation. As a matter of fact, I have such reservations myself: in Williamson and Mahar (1998) interest rate liberalization is identified as merely one of six dimensions of financial liberalization. Moreover, Stiglitz (1994) has argued that interest rate liberalization should come toward the end of the process of financial liberalization, inasmuch as a ceiling on the deposit interest rate (equal to the Treasury bill rate, he suggests) might provide a constraint on gambling for redemption. I find this argument persuasive and long ago changed my description of the fourth element of the Washington Consensus to financial liberalization. More recently Stiglitz (1998) has expressed a much more basic objection to financial liberalization, arguing that the success of some East Asian countries stemmed importantly from their policy of directing credit to particular industries rather than allowing the market to determine the allocation of credit. That argument is highly contentious, especially in the aftermath of the East Asian economic crisis of 1997–98.

My fifth choice—a competitive exchange rate—was not, I have concluded, an accurate report of Washington opinion. I suspect that by 1989 a majority of economists, in Washington as elsewhere, were already in favor of either firmly fixed or freely floating exchange rates and hostile to the sort of intermediate regime that in my judgment gives the best promise of maintaining a competitive exchange rate in the medium term. (My own preference remains an intermediate regime of limited flexibility, provided that excludes an old-fashioned adjustable peg, even if such a regime is more likely to spawn speculative pressures than a floating rate.) But note that the East Asian countries did by and large achieve and maintain competitive exchange rates, at least before about 1996 (and even after 1996 only Thailand failed to do so).[2]

My sixth reform was trade liberalization. Here I see little reason to doubt that I reported accurately on opinions in the international financial institutions and the central economic agencies of the U.S. government (although parts of Congress and the Department of Commerce are not noted for their dedication to liberal trade). But this is another area where critics can rightly claim that the policies that nurtured

the East Asian miracle were, at least in some countries, at odds with the policies endorsed in the Washington Consensus. Much the same is true of foreign direct investment, except that the East Asian economies were less hostile to a policy of openness; only the Republic of Korea rejected most foreign direct investment during the years of the miracle.

Privatization commanded a lot of support in Washington, where it had been put on the international agenda by James Baker when he was secretary of the U.S. Treasury, in his speech to the World Bank–International Monetary Fund Annual Meetings in Seoul in 1985. Privatization was controversial in much of the rest of the world, where one's attitude to public versus private ownership had long been the litmus test for qualifying as left-wing or right-wing. Deregulation was rather less politically polarizing: it had been initiated by the centrist Carter administration in the United States, rather than by the right-wing Thatcher government that pioneered privatization in the United Kingdom. Deregulation, however, was not a policy that reverberated in East Asia, where the industrial policies pursued in some (though not all) countries ran very much in the opposite direction. The notion of the importance of secure property rights had come both from Chicago's law and economics school and from the work of Hernando de Soto in Peru. The concept was presumably offensive to those who resisted the advance of the market economy, but this breed was extinct in Washington by 1989 (if, indeed, it had ever existed there). My impression is that the institution of private property was somewhat more securely entrenched in East Asia than in most of the rest of the developing world.

So much for the content of my version of the Washington Consensus. What inspired it? In an immediate sense, it originated from an attempt to answer a question posed to me by Hans Singer during a seminar at the Institute for Development Studies: what were these "sensible" policies that were being pursued in Latin America (and that I was arguing justified approval of the Brady Plan to provide these countries with debt relief)? In a more profound sense, my effort was an attempt to distill which of the policy initiatives that had emanated from Washington during the years of conservative ideology had won inclusion in the intellectual mainstream rather than being cast aside once Ronald Reagan was no longer on the political scene.[3] Taking an even longer perspective, my version of the Washington Consensus can be seen as an attempt to summarize the policies that were widely viewed as supportive of development at the end of the two decades when economists had become convinced that the key to rapid economic development lay not in a country's natural resources or even in its physical or human capital but, rather, in the set of economic policies that it pursued.

Let me emphasize that the Washington Consensus as I conceived it was in principle geographically and historically specific, a lowest common denominator of the reforms that I judged "Washington" could agree were needed in Latin America as of 1989. But in practice there would probably not have been a lot of difference if I had

undertaken a similar exercise for Africa or Asia, and that still seemed to be the case when I revisited the topic (with regard to Latin America) in 1996 (Williamson 1997). This doubtless made it easier for some to interpret the Washington Consensus as a policy manifesto that its adherents supposedly believed to be valid for all places and at all times.

Current Usage

The following is a selection of recent definitions of the Washington Consensus that I happened to stumble across. (I have undertaken no bibliographic research to compile this list.)

"A die-hard liberalization advocate (or a Washington-consensus believer). . . . " (Ito 1999)

". . . the self-confident advice of the 'Washington consensus'—free-up trade, practice sound money, and go home early. . . ." (Vines 1999)

" . . . the Washington Consensus: policy prescriptions based on free market principles and monetary discipline." (Hamada 1998)

"The Washington Consensus had the following message: 'Liberalize as much as you can, privatize as fast as you can, and be tough in monetary and fiscal matters.'" (Kolodko 1998)

"The bashing of the state that characterized the policy thrust of the Washington Consensus. . . ." (United Nations 1998)

"This new imperialism, codified in the 'Washington Consensus'. . . ." (Alam 1999)

"The Brazilian crisis has reignited the debate over the so-called Washington Consensus on the creation of a laissez-faire global economy." (Rajan 1999)

In none of these examples is my phrase used in the sense that I originally intended. On the contrary, when I coined the term in 1989, the market fundamentalism of Reagan's first term had already been superseded by the return of rational economic policymaking, and one could discern which ideas were going to survive and which were not (monetary discipline but not monetarism; tax reform but not tax-slashing; trade liberalization but maybe not complete freedom of capital movements; deregulation of entry and exit barriers but not the suppression of regulations designed to protect the environment).

How is it that a term intended to describe a technocratic policy agenda that survived the demise of Reaganomics came to be used to describe an ideology embracing the most extreme version of Reaganomics? The closest I can come to understanding this is to note that my version of the Washington Consensus did indeed focus principally on policy reforms that reduced the role of government, such as privatization and the liberalization of trade, finance, foreign direct investment, and entry and exit. It did this because the orthodoxy of the generation whose ideas were embodied in

the practices being challenged in 1989 had been much more statist than was by then regarded as advisable, and hence the policy reforms that were needed at that time were all in the direction of liberalization. This need for liberalization did not necessarily imply a swing to the opposite extreme of market fundamentalism and a minimalist role for government, but such boring possibilities were repressed in the ideological debates of the 1990s. For it is certainly true that the Washington Consensus came to be used to describe an ideological position, a development that Naim (2000) argues resulted from the world's acute need for a new ideology to provide a focus for debate in place of the god that had failed. My qualifications about the Washington Consensus being an agenda for a specific part of the world at a particular moment of history were quickly forgotten, as the search for a new ideology, to endorse or to hate, was perceived to have succeeded. Ravi Kanbur argues that the staffs of the Bretton Woods institutions perceived themselves as storming the citadels of statism, which led them as a negotiating ploy to demand more in the way of liberalizing reforms than they really expected to achieve—a tactic that led citizens in the World Bank's client countries to identify these institutions with something closer to market fundamentalism than the institutions really believed in.

The term's use as a synonym for market fundamentalism appears to be the dominant, but not the only, current usage. Many Bank staff members, including those who wrote *Beyond the Washington Consensus: Institutions Matter* (Burki and Perry 1998), still use the term in the way that I intended, and I think most of them would endorse the reform agenda to which I had applied the term as a reasonably accurate and appropriate summary of what the Bank and other agencies concerned with the promotion of development were, and should have been, advising countries to do.

Joseph Stiglitz, formerly the World Bank's chief economist, recently used the term in the alternative, neoliberal, sense (1999b). This at least makes it clear that he was not attacking his colleagues when he spoke of reviewing "the major ways in which . . . the 'Washington Consensus' doctrines of transition, failed. . ." (Stiglitz 1999a:4). He proceeded to question the priority given to rapid privatization and the lack of attention to establishing competition or building social and organizational capital, and later he spoke of "the standard form of voucher privatization promoted by the Washington Consensus. . . ." I am not aware that Washington has ever displayed any particular preference for voucher privatization; certainly this was not a theme of the 1996 *World Development Report* (World Bank 1996), which dealt with the transition. I agree with Stiglitz on the substantive questions he raises: one can put too much emphasis on rapid privatization, and it is more important to do it right than to do it quickly; I agree that the great merit of privatization is that it can be used to further competition; I am skeptical about voucher privatization; and I think I agree about the importance of social and organizational capital, if I understand what the words mean. (I would describe them as social cohesion and good institutions, respectively.) What I do not understand is what is gained by describing these sensible

ideas as refuting a doctrine described by a term that many people in the Bank regard as providing a useful summary of the advice the Bank dispenses.

Do Washington Consensus Policies Promote Poverty Reduction?

The answer, quite obviously, depends on which interpretation of the Washington Consensus one is referring to. The popular, or populist, interpretation of the Washington Consensus, meaning market fundamentalism or neoliberalism, refers to laissez-faire Reaganomics—let's bash the state, the markets will resolve everything. I would not subscribe to the view that such policies offer an effective agenda for reducing poverty. We know that poverty reduction demands efforts to build the human capital of the poor, but the populist interpretation fails to address that issue. We know that an active policy to supervise financial institutions is needed if financial liberalization is not to lead to financial collapse, which invariably ends up using tax revenues to write off bank loans that were made to the relatively rich. And some measure of income redistribution would be recommended by any policy that was primarily directed at reducing poverty rather than simply maximizing growth, but market fundamentalists rule out all income redistribution as plunder.

A plausible alternative concept would be that the Washington Consensus consists of the set of policies endorsed by the principal economic institutions located in Washington: the U.S. Treasury, the Federal Reserve Board, the International Monetary Fund, and the World Bank. I would argue that the policies these institutions advocated in the 1990s were inimical to the cause of poverty reduction in emerging markets in at least one respect: their advocacy of capital account liberalization. This was, in my view, the main cause of the contagion that caused the East Asian crisis to spread beyond Thailand and that resulted in a tragic interruption of the poverty reduction those countries had achieved (Williamson 1999). (I did not include full capital account liberalization in my version of the Washington Consensus because I did not believe it commanded a consensus, if only because I could not believe I was the only person in Washington who feared that capital account liberalization could precipitate a tragedy such as that which occurred in East Asia.)

My version of the Washington Consensus began with the proposition that the inflation caused by lack of fiscal discipline is bad for income distribution. The second reform specifically involved redirecting public expenditure toward primary health and education, that is, toward building the human capital of the poor. Tax reform can be distributionally neutral or even progressive. A competitive exchange rate is key to nurturing export-led and crisis-free growth and is hence in the general interest, including that of the poor. Trade liberalization, certainly in low-income, resource-poor countries, tends to be pro-poor because it increases the demand for unskilled

labor and decreases the subsidies directed to import-competing industries that use large volumes of capital and employ small numbers of workers, many of them highly skilled. Foreign direct investment helps raise growth and spread technology, provided that import protection is not excessive, so that the case of immiserizing growth does not arise (Brecher and Diaz-Alejandro 1977).[4] The impact of privatization depends very much on how it is done: the sort of insider-voucher privatization that occurred in Russia allows the plunder of state assets for the benefit of an elite, but a well-conducted privatization with competitive bidding can raise efficiency and improve the public finances with benefits to all, including the poor. Deregulation in general involves the dismantling of barriers that protect privileged elites (even if some of them, like trade unionists, have difficulty thinking of themselves as an elite), and hence there is a strong presumption that it will be pro-poor. Private property rights are certainly a defense primarily for those who have private property, but the improvement of such rights is nonetheless very likely to be pro-poor because these are the people who find themselves unable to defend their property when property rights are ill-defined (for example, Hernando de Soto's squatters on the periphery of Lima).

I have omitted one of the ten reforms from the preceding list: financial and interest rate liberalization. This is the primary focus of Stiglitz's criticisms when he refers to something that I can recognize as akin to my version of the Washington Consensus. I have realized for some time (see Williamson 1996) that my first formulation was flawed in that it neglected financial supervision, without which financial liberalization seems all too likely to lead to improper lending and eventually to a crisis that requires the taxpayers to pick up the losses from making bad loans (Williamson and Mahar 1998). But should economists therefore endorse the view that directed lending as pursued in some—though not all—East Asian countries is pro-growth and thus ultimately pro-poor? On this issue, at least, I would have thought that the East Asian crisis, especially in Korea, should have tempered economists' enthusiasm for the practice. The high debt-equity ratios that resulted from directed lending were certainly among the causes of the financial fragility that deepened the impact of the crisis.

Thus most of the reforms embodied in my version of the Washington Consensus are at least potentially pro-poor. In some cases this conclusion is sensitive to the way in which reform is implemented: that is certainly true of tax reform, privatization, and, above all, financial liberalization. But I see no reason why the World Bank should back away from endorsing my version of the Washington Consensus in view of its reaffirmation of poverty reduction as its overarching mission. That is not to claim that the Washington Consensus, in any version, constituted a policy manifesto adequate for addressing poverty. My version quite consciously eschewed redistributive policies, taking the view that Washington had not reached a consensus on their

desirability. But time has moved on, and we are now looking to *World Development Report 2000/01* for an outline of the policies needed to supplement my version of the Washington Consensus in a world that takes poverty reduction seriously.

The Semantic Dilemma

One can react to the semantic dilemma posed by the different definitions currently in use in three possible ways. Consider these alternatives:

- *Insist on the original usage.* Insist that my version of the Washington Consensus is the only correct and legitimate interpretation, as a corollary of which the term will (with the qualifications noted above) be recognized as pro-poor. This alternative strikes me as both presumptuous and unrealistic: once a term has escaped into the public domain, one cannot dictate the reestablishment of a common usage. The likely result would be a perpetuation of the public confusion that I am attempting to address.
- *Abandon the term.* Refuse to debate in the terms that have been so compromised by the widespread adoption of the "populist" definition. I cannot imagine that this approach would end the populist use of the term; it would simply be a cop-out.
- *Endorse a post–Washington Consensus.* A more promising strategy has been adopted at least twice within the Bank. In 1998 the Latin America Regional Office of the World Bank issued a policy document that favored going beyond the Washington Consensus (Burki and Perry 1998). Stiglitz did almost the same, semantically at least, in urging a post–Washington Consensus in his lecture to the World Institute for Development Economics Research in January 1998.

When I first came across this approach, I thought it implied that the reforms included in the Washington Consensus were necessary but not sufficient for promoting development, an idea that seemed eminently reasonable. Clearly the Bank today would want to go further and endorse a wider array of antipoverty instruments than was able to command a consensus in 1989, when the most I thought I could legitimately include was the promotion of public expenditure on primary health and education.[5]

In their book, Burki and Perry (1998) explicitly refer to my version of the Washington Consensus and assert that the widespread implementation of the "first-generation" reforms it prescribed was paying off in Latin America in resumed growth and an end to high inflation. They noted that the reforms had not been equally effective in reducing poverty and inequality, which they argued demonstrated a

"need to focus on improving the quality of investments in human development, promoting the development of sound and efficient financial markets, enhancing the legal and regulatory environments (in particular, deregulating labor markets and improving regulations for private investment in infrastructure and social services), [and] improving the quality of the public sector (including the judiciary) . . ." (p. 4). This is an agenda dominated by institutional reform, which is indeed what has become known in Latin America as the second-generation reform agenda (Naim 1995).

It is not equally obvious why Stiglitz would want to propagate a post–Washington Consensus that implied endorsing and extending the original version, given his interpretation of what was included in it. In fact, the Stiglitz version of a post–Washington Consensus does not endorse any version of the original. He is advocating a policy package that is intended to supersede the Washington Consensus altogether. His new policy package is asserted to differ from the original in two dimensions.

First, he argues that the implicit policy objective underlying the Washington Consensus is inadequate. In addition to pursuing economic growth, the objectives should include "sustainable development, egalitarian development, and democratic development." In other words, he believes that policy objectives should include the state of the environment, income distribution, and democracy, as well as per capita gross national product. I find those objectives much more congenial than a single-minded preoccupation with economic growth, although I am not sure that the World Bank could formally endorse the pursuit of democracy (its Articles do, after all, forbid its involvement in politics).[6] Second, in addition to expanding the objectives, Stiglitz argues that it is necessary to pursue "sound financial regulation, competition policy, and policies to facilitate the transfer of technology and transparency" to make markets work in a way that will support development.

I have a somewhat different view of what should be added to the Washington Consensus to make it a policy manifesto supportive of egalitarian, environmentally sensitive development. I agree that financial regulation (prudential supervision) is crucial and that transparency is a useful complement to supervision in achieving appropriate conduct of financial institutions. Moreover, competition is a natural complement to deregulation in promoting a well-functioning market economy (although a liberal import regime is the most effective competition policy in tradables, as Srinivasan argues in his comment in this volume). I would not have included technology transfer in such a manifesto, although I would have no objection to including institutional changes that seemed likely to promote technology transfer if I were reasonably confident that I knew what these changes were (besides accepting foreign direct investment). Similarly, I would consider it desirable to include policies focused on improved environmental conditions, although I am not sure that I would know how to select policy measures at a comparable level of generality to my 10

original points. But my emphasis would have been different; I would have focused much more generally on institutions. To explain why, let me offer a brief history of postwar development thinking.

In the first wave of theorizing about economic development, from the 1940s to the early 1960s, economists saw the accumulation of physical capital as the key to development (as reflected in the Harrod-Domar model, the Lewis model, and the two-gap model). The second phase recognized that human capital provided another and more inelastic constraint on development, a constraint that explained why Europe and Japan had recovered from World War II so rapidly, when growth in developing countries had been lagging despite the adoption of development policies and the beginning of large-scale aid. The third phase, which started about 1970 with the work of Little, Scitovsky, and Scott (1970) and Balassa (1970), emphasized that the policy environment influenced the level and dominated the productivity of investment. The Washington Consensus attempted to summarize the outcome of this debate on the policies that were conducive to economic development. The major advance of the 1990s stemmed from recognition that the central task of the transition from communist to market-based economies involved building the institutional infrastructure of a market economy. This realization was complemented by a growing recognition that bad institutions can sabotage good policies. This viewpoint was reflected in Stiglitz's (1999a) remarks on the transition, in Naim's (1995) work on supplementing the Washington Consensus, in Burki and Perry (1998), in the *World Development Reports* of 1997 and 1998, and in the World Bank's decision to launch a crusade against corruption.

What should one make of the idea of launching a post–Washington Consensus? I would not be happy at such a move if it were interpreted to imply a rejection of "the" Washington Consensus, although I would have no problem if it involved rejection of the populist, or market-fundamentalist, version. But it seems a somewhat odd crusade. The time of the original consensus, 1989, was an unusual period in that the ideological battles of the Reagan era, not to mention the cold war battle between capitalism and communism, were passing into history, leaving in their wake an unusually wide measure of agreement that several rather basic ideas of good economics were not only desirable but of key importance in the current policy agenda of at least one region—Latin America. Currently, there is no similar coalescing of views, certainly not on the wider agenda that Stiglitz has laid out. (Consensus on egalitarianism? With aid fatigue threatening the future of the International Development Association? On environmental sustainability? In a world where the U.S. Senate refuses even to consider ratifying the Kyoto Protocol?) I agree, rather, with Tim Geithner (1999:8): "I don't think anyone believes there is some universal model that can or should be imposed on the world—Washington consensus, post Washington consensus, or not."

Resolving the Dilemma

Let me conclude by laying out my own ideas on how to resolve the dilemma.

- There is little merit in attacking abstract, undefined concepts that are interpreted to mean whatever the author momentarily decides they mean. It is better to spell out those concepts that are being criticized and debate policies on the basis of their merits.
- The World Bank should recognize that the term Washington Consensus has been used in very different ways. One summarizes policies that are pro-poor; another describes a policy stance that offers the poor very little and warrants no support.
- It is appropriate to go beyond the Washington Consensus by emphasizing the importance of the institutional dimension as well as of the sort of policies embodied in the original version of the Washington Consensus—policies that will promote an equitable distribution of income as well as a rapid growth of income.
- The hopeless quest to identify a consensus where there is none should be abandoned in favor of a debate on the policy changes needed to achieve a rounded set of objectives encompassing at least the level, growth, and distribution of income, as well as preservation of a decent environment.

The Bank will do the cause of economic development a great service if it can frame future debate in these terms. Admittedly my suggestions do not answer the pleas for a new ideology that would more adequately reflect the goals of the multilateral development banks and that might thus increase the chance of establishing local ownership of the sort of economic policy stance conducive to rapid and equitable growth. Let me plead in defense that I am not a suitable person to launch an ideology, inasmuch as Naim (2000) characterizes an ideology as a thought-economizing device and I actually believe that thinking is more desirable than economizing on thought.

Notes

John Williamson is senior fellow at the Institute for International Economics. This article was written as a background paper for *World Development Report 2000/01*. The author is indebted to the participants in a session at which an early version of the paper was discussed, notably Ravi Kanbur and Moisés Naím.

1. This intellectual convergence was the result of the collapse of communism, which resulted not from machinations of the Bretton Woods institutions, or even of the U.S. Central Intelligence Agency, but because socialism does not work except in a simple economy, and even then it seems to have worked reasonably well only when large numbers of people were inspired with revolutionary zeal.

2. Exchange rate policy is the one topic on which I have a serious difference of view with T. N. Srinivasan's comment that accompanies this paper. The term "competitive exchange rate" originated with Bela Balassa and signifies a rate that is either at, or undervalued relative to, its long-run equilibrium. I do not regard measuring the latter as an exercise in futility; see Hinkle and Montiel (1999) for evidence that other people in the Bank do not either. I dissent from the consensus Srinivasan proclaims that holds that only currency boards and freely floating rates offer viable regimes. For further details, see Williamson (forthcoming).

3. In trying to identify policies from the Reagan-Thatcher era that had not won consensus support, I wrote in 1996: "it [the Washington Consensus] did not declare that the only legitimate way to restore fiscal discipline was to slash government expenditure; it did not identify fiscal discipline with a balanced budget; it did not call for overall tax cuts; it did not treat as plunder the taxes raised to redistribute income; it did not say that exchange rates had to be either firmly fixed or freely floating; it did not call for the proscription of capital controls; it did not advocate competitive moneys or argue that the money supply should grow at a fixed rate" (Williamson 1997:50).

4. Growth of output of a heavily protected product can immiserize a country if the resources used in production exceed the social value of the output.

5. However, in commenting on my paper, Stanley Fischer (then the Bank's chief economist) argued that I could and should have gone further: "Emphasis on poverty reduction has increased in recent years and will continue to do so. [A good forecast.] The concern with poverty reduction goes beyond the belief that economic growth will reduce poverty, to the view that targeted food subsidies as well as the medical and educational programs to which Williamson refers, can reduce the number of poor people . . . and should be used for that purpose" (Williamson 1990:27).

6. Some people might wish to add nation-building to the noneconomic objectives to be pursued by development policy (as was common in the 1960s).

References

The word "processed" refers to informally reproduced works that may not be commonly available through library systems.

Alam, M. Shahid. 1999. "Does Sovereignty Matter for Economic Growth?" In John Adams and Francesco Pigliaru, eds., *Economic Growth and Change.* Cheltenham, U.K.: Edward Elgar.

Balassa, Bela. 1970. "Growth Strategies in Semi-Industrial Countries." *Quarterly Journal of Economics* 84(1):24–47.

Brecher, Richard, and Carlos Diaz-Alejandro. 1977. "Tariffs, Foreign Capital, and Immiserizing Growth." *Journal of International Economics* 7(4):317–22.

Burki, Javed, and Guillermo E. Perry. 1998. *Beyond the Washington Consensus: Institutions Matter.* Washington, D.C.: World Bank.

Geithner, Timothy F. 1999. "The World Bank and the Frontier of Development Challenges." Remarks delivered to a conference on "Reinventing the World Bank: Challenges and Opportunities for the 21st Century," Northwestern University, May 14. Processed.

Hamada, Koichi. 1998. "IMF Special: Keeping Alive the Asian Monetary Fund." *Capital Trends* 3(September).

Hinkle, Lawrence E., and Peter J. Montiel, eds. 1999. *Exchange Rate Misalignment: Concepts and Measurement for Developing Countries.* New York: Oxford University Press.

Ito, Takatoshi. 1999. "The Role of IMF Advice." Paper presented to International Monetary Fund conference on "Key Issues in Reform of the International Monetary and Financial System," Washington, D.C., May 29. Processed.

Kolodko, Gregorz. 1998. *Transition.* World Bank Development Economic Research Group newsletter, June.

Little, Ian, Tibor Scitovsky, and Maurice Scott. 1970. *The Structure of Protection in Developing Countries: A Comparative Study.* Baltimore, Md.: Johns Hopkins University Press.

McCloskey, Deirdre N. 1998. *The Rhetoric of Economics.* 2d ed. Madison, Wis.: University of Wisconsin Press.

Naím, Moisés. 1995. "Latin America: The Morning After." *Foreign Affairs* 74(July–August):45–61.

———. 2000. "Washington Consensus or Washington Confusion?" *Foreign Policy* (spring):86–103.

Rajan, Ramkishen S. 1999. "The Brazilian and Other Currency Crises of the 1990s." *Claremont Policy Brief.* Claremont College, Claremont, Calif. May.

Soros, George. 1998. *The Crisis of Global Capitalism.* New York: Public Affairs.

Stewart, Frances. 1997. "Williamson and the Washington Consensus Revisited." In Louis Emmerij, ed., *Economic and Social Development into the XXI Century.* Washington, D.C.: Inter-American Development Bank.

Stiglitz, Joseph E. 1994. "The Role of the State in Financial Markets." In Michael Bruno and Boris Pleskovic, eds., *Proceedings of the World Bank Conference on Development Economics 1993.* Washington, D.C.: World Bank.

———. 1998. "More Instruments and Broader Goals: Moving toward the Post–Washington Consensus." United Nations University/World Institute for Development Economics Research, Helsinki.

———. 1999a. "Whither Reform? Ten Years of the Transition." In Boris Plesovic and Joseph E. Stiglitz, eds., *World Bank Annual Conference on Development Economics.* Washington, D.C.: World Bank.

———. 1999b. "The World Bank at the Millennium." *Economic Journal* 109(459):F577–97.

United Nations. 1998. *Annual Report of the United Nations University.* Tokyo.

Vines, David. 1999. In an obituary for Susan Strange. Newsletter no 9. Economic and Social Research Council Programme on Global Economic Institutions, Swindon, U.K. Processed.

Waelbroeck, Jean. 1998. "Half a Century of Development Economics: A Review Based on the *Handbook of Development Economics.*" *The World Bank Economic Review* 12(May):323–52.

Williamson, John. 1990. "What Washington Means by Policy Reform." In John Williamson, ed., *Latin American Adjustment: How Much Has Happened?* Washington, D.C.: Institute for International Economics.

———. 1996. "Are the Latin American Reforms Sustainable?" In Hermann Sautter and Rolf Schinke, eds., *Stabilization and Reforms in Latin America: Where Do We Stand?* Frankfurt: Vervuert Verlag.

———. 1997. "The Washington Consensus Revisited." In Louis Emmerij, ed., *Economic and Social Development into the XXI Century.* Washington, D.C.: Inter-American Development Bank.

———. 1999. "Implications of the East Asian Crisis for Debt Management." In A. Vasudevan, ed., *External Debt Management: Issues, Lessons, and Preventive Measures.* Mumbai: Reserve Bank of India.

———. Forthcoming. *Exchange Rate Regimes for Emerging Economies: Reviving the Intermediate Option.* Washington, D.C.: Institute for International Economics.

Williamson, John, and Molly Mahar. 1998. "A Survey of Financial Liberalization." Princeton Essays in International Finance 211. Princeton University, Princeton, N.J.

World Bank. 1993. *The East Asian Miracle: Economic Growth and Public Policy.* New York. Oxford University Press.

———. 1996. *World Development Report: From Plan to Market.* New York: Oxford University Press.

The Washington Consensus a Decade Later: Ideology and the Art and Science of Policy Advice

T. N. Srinivasan

In 1990 Williamson coined the term "Washington Consensus" to describe "the lowest common denominator of policy advice being addressed by the Washington institutions to Latin American countries as of 1989." He now protests in an article in this volume that the phrase has "become a synonym for 'neoliberalism' or what George Soros (1998) has called 'market fundamentalism'. . . . " This development should have caused no surprise, given the visceral hatred in many parts of the world for free markets. Because they view the U.S. government and its "lackeys"—the World Bank, the International Monetary Fund, and to a lesser extent, the Inter-American Development Bank—as the chief advocates of free markets, these opponents of free markets would have attacked anything called the "Washington Consensus." Had he instead called it the "Williamson synthesis," perhaps the package would have been a less inviting target for attack, although one cannot be sure: given the author's association with the hated Washington institutions, he could not have credibly absolved them from responsibility for the synthesis.

More than a decade has passed since Williamson proffered the consensus package. It is tempting (following the advice of the late Senator George Aiken of Vermont, who suggested that the United States simply declare victory and withdraw from the then-raging Vietnam War) to announce that the policy prescriptions contained in the consensus package have been successful and are no longer contentious! Even though there is considerable evidence of the success of many elements of the package where appropriately implemented, unfortunately the top echelon of at least one of the institutions celebrated in the consensus—the World Bank—seems to be misinformed and confused about the contents of the policy package. Others doubt the appropriateness of some of the recommended policies and argue that other elements of the package never commanded a consensus. It is useful, therefore, to reexamine

The World Bank Research Observer, vol. 15, no. 2 (August 2000), pp. 265–70.

the package for its contemporary relevance, taking into account the experience of the 1990s.

At the outset, let me say that I regard the policy package less as a consensus and more as a reflection of the author's synthesis of lessons from four decades of development experience. Indeed, Williamson's phrase "common denominator of policy advice" can mean only one thing: that the advice is based on policy implications that emerged not only from a number of studies of certain Latin American countries but also, more generally, from experience in many developing countries. In this sense, trade liberalization as a sensible policy was supported by studies published in the 1970s by the Organisation for Economic Co-operation and Development (Little, Scitovsky, and Scott 1970), the National Bureau of Economic Research (Bhagwati 1978; Krueger 1978), and the World Bank (Balassa 1970). (A later World Bank study, Papageorgiou, Michaely, and Choksi 1991, lent similar support.) But only the utterly naive would interpret this advice as unqualified in the sense of being appropriate regardless of the specific circumstances of an economy or the time horizon involved. Yet the studies covered many countries with varying socio-political-economic institutions and many different time periods, and almost all of them still found possible benefits from trade liberalization—a fact that testifies to the robustness of the policy advice. Be that as it may, it is common knowledge among economists that the response to any policy change, such as trade liberalization, that operates through price incentives depends both on nonprice factors and on the time horizon. For example, if domestic supply constraints (other than price received) are severe in the short and medium run, removing all price distortions would have only a limited favorable response. And, to the extent that tax revenues are largely derived from trade taxes, the government may be constrained as to how far it can liberalize trade by reducing trade taxes without compromising fiscal discipline. This constraint may not be binding, however, if the actual applied tariffs exceed revenue-maximizing levels, as is often claimed. In any case, one can enumerate many nuances and caveats to each of the recommended policies in the package. But only an ideologue or the utterly ignorant would conclude that because caveats apply, any attempt to change the status quo through the implementation of the recommended policies is undesirable.

Similarly, the advice to liberalize interest rates was a common conclusion from several studies of financial repression in developing countries. Williamson now concedes that it could be costly to liberalize interest rates before other elements of financial liberalization, such as prudential supervision of banks by capable and knowledgeable central bank authorities, are in place. Although this point is obviously valid, the scope of its applicability is arguable. For example, in many developing countries, most banks are publicly owned, and whether they would or could gamble for redemption if deposit interest rates were to be liberalized is open to question. At any

rate, as in the case of trade liberalization, the case for interest rate liberalization is likely to remain intact for many developing countries.

The advice on fiscal discipline was also based on the experience of many developing countries, particularly Latin American countries, that had previously undergone episodes of hyperinflation and stop-and-go sequences of stagnation and growth. India's experience in the 1980s, when it abandoned fiscal discipline to run deficits financed by costly borrowing at home and abroad, is instructive. The spurt in growth following the reckless fiscal expansionism proved unsustainable and ended inevitably in a macroeconomic-cum-balance of payments crisis in 1991. The fact that India did not experience Latin-style inflation before or after the abandonment of fiscal discipline is beside the point: because most of India's poor workers, particularly those in agriculture and informal service sectors, are not protected against inflation, even moderate inflation by Latin standards has serious consequences for the welfare of the poor in India.

The advice to redirect public expenditure toward health care, primary education, and infrastructure has long been part of conventional wisdom. A point that is not part of the conventional wisdom (and one that leading advocates of redirection such as Amartya Sen and his acolytes do not emphasize) is that failure to liberalize trade, to privatize inefficient public enterprises of dubious social value (such as airlines, hotels, and steel mills), and to reform the tax system eats away public resources that could otherwise be directed to the social sectors. Thus trade liberalization, privatization, and tax reform, which constitute three of the ten policies in the consensus package, are important not only in and of themselves but also because they make more resources available for social sectors.

The package called for "a competitive exchange rate," an unfortunate choice of words. An exchange rate is the price of one currency in terms of another currency or a basket of currencies. The word "competitive" used in the context of an exchange rate evokes painful memories of competitive devaluations by many countries in the interwar era. Under the classic Bretton Woods system of fixed exchange rates, one could interpret an "uncompetitive" exchange rate to mean one that is overvalued relative to its long-run equilibrium value. The operational significance of the interpretation is vastly diminished, however, by the fact that the long-run equilibrium rate is hard to define, that it was not defined in the Articles of Agreement of the International Monetary Fund, and that it is not simple to compute from available data in any case. Perhaps by a competitive exchange rate Williamson meant an undervalued exchange rate and that in fact Japan and the East Asian economies following Japan maintained competitive exchange rates in this sense.

Edwards and Savastano (1999) recently surveyed the empirical studies on exchange rate regimes in developing countries. They identified two camps: one group ascribes a key role to the exchange rate as a nominal anchor; the other stresses the perils of

relying on an asset price and therefore sees the exchange rate as a nominal anchor in a world of integrated global capital markets. In their view

> . . . the differences between the two broad camps identified lie in their differing views regarding three key features of exchange rate policy in a context of high capital mobility: (1) the scope for (and effectiveness of) sterilized and unsterilized intervention as a means for attaining (and preserving) a degree of nominal exchange rate stability; (2) the costs that "excessive" fluctuations of the nominal exchange rate may impose on the economy's performance; and (3) the *time dimension* of their analysis—i.e., the horizon over which monetary policy, the exchange rate, capital flows and the rest of the economy are assumed to interplay. All of these are empirical issues for which little, if anything, is known for the case of developing countries—not even for the relatively advanced ones. (Edwards and Savastano 1999:22)

Since the Asian financial crisis, a consensus seems to be emerging that the only two viable exchange rate regimes are either a system of rigidly fixed exchange rates implemented through a currency board arrangement or its opposite, a regime of freely floating exchange rates. Williamson, however, now prefers an intermediate regime between the two with limited flexibility. He does not explain why and how the limited flexibility of the regime could be credibly signaled to distinguish it, on the one hand, from the old-fashioned and now discredited crawling peg and, on the other hand, from a regime of transition to a free float. The advice to keep the exchange rate "competitive" has no operational content even if a competitive exchange rate could be defined in conceptual terms. Of course, allowing the rate to float freely obviates this problem.

The recommendation to liberalize flows of foreign direct investment grew in part out of the need to have capital inflows that did not create debt and in part from the desire for other benefits such as technology transfers, which were believed to be associated with such investments. Such investment flows are not likely to be as volatile as short-term capital flows because a decision to invest in another country is probably based on the long-term fundamentals of the recipient economy. Even as late as 1989, private capital flows, including foreign direct investment, had not accelerated as much as they did in the 1990s. The financial crises starting with Mexico in December 1994 and ending with the most recent one in Brazil in 1998 have exposed several weaknesses in the domestic financial sectors of developing countries and in the international financial architecture. The crises have, if anything, reinforced the advice to liberalize foreign direct investment flows.

The advice to deregulate, in the sense of abolishing barriers to entry and, equally important, to exit, continues to be pertinent because it is based on age-old and proven virtues of competition. Indeed, benefits from removing price distortions will be limited

if firms not strong enough to compete in an undistorted market are not allowed to exit and if more efficient firms are denied entry.

The obverse side of deregulation is the need to ensure competition to firms that are privatized. For internationally traded goods and services (except possibly in wide-body passenger jet aircraft!), the world market is far larger than the minimum efficient scale of production, so that opening to trade is adequate to generate competition. In nontraded sectors where considerations of scale economies or network externalities preclude the possibility of more than a handful of firms operating on an efficient scale, a regulatory authority needs to be established and an appropriate set of regulations promulgated so that the few private firms operate in a socially desirable manner. Interestingly, technological developments have vastly eroded scale economies in electricity generation, telecommunications, and other nontraded goods and services that were once deemed natural monopolies.

In 1990, when Williamson published his consensus, regulation and privatization in developing countries were scarce, and there was none in Russia and the Eastern European countries, which were still part of the disintegrating Soviet Union. In the decade since then, privatization has taken off, particularly in the transition economies. It is fair to say, however, that the state of knowledge about appropriate mechanisms for privatization and regulatory frameworks (particularly for private financial intermediaries) is still in a state of flux. Nonetheless, the advice to privatize remains relevant, with a cautionary note about the need to ensure sufficient competition for privatized enterprises.

My brief evaluation of the Washington Consensus 10 years after its promulgation strongly suggests that its policy advice remains largely intact. In general, sound policy advice, while undoubtedly based on received theory and empirical evidence, necessarily has to involve judgment. Economic theorizing involves simplification and abstraction of complex reality; econometric analysis of empirical evidence often imposes restrictions on theory such as, for example, specific functional forms for utility and production functions, the nature of heterogeneity among firms and consumers, and distributional assumptions about stochastic terms. Simple and abstract theory (which in its "second-best" version could amount to saying that almost anything is possible!) and its highly restrictive econometric specification cannot deliver policy conclusions that can be directly applied to the situation of any given economy at a particular time. To advise on policy requires sound judgment on the part of the advisor—judgment that goes beyond findings for theoretical and econometric models. Such judgment will incorporate knowledge about history, particularly economic history, and about the specific socio-political-cultural features and institutions of the country involved. For example, cross-country regressions, even if they are not mindless, cannot deliver policy conclusions about the desirability of trade liberalization or capital inflows or about the effect of openness on growth. At best, such regressions are an efficient means for discovering patterns in the data from which capable re-

searchers can draw inferences after bringing to bear their knowledge about the economies involved while firmly eschewing attribution of causality. To say that judgment based on the specifics of the case is needed is not to argue either that there can be no robust policy conclusions of wide applicability or that only discretion, rather than rules, should govern policy choice. What it means is that formal analysis has to be supplemented by informally allowing for factors that by necessity have been excluded from the formal analysis. There is art as well as science in policy advising! An honest policy advisor will clearly indicate where his judgment enters and where theory and econometrics stop. It is not hard to isolate the theoretical and econometric bases and the astute judgment of Williamson in the Washington Consensus.

Note

T. N. Srinivasan is Samuel C. Park Jr. Professor of Economics at Yale University.

References

The word "processed" describes informally reproduced works that may not be commonly available through library systems.

Balassa, Bela. 1970. "Growth Strategies in Semi-Industrial Countries." *Quarterly Journal of Economics* 84(1):24–47.

Bhagwati, Jagdish. 1978. *Foreign Trade Regimes and Economic Development: Anatomy and Consequences of Exchange Control Regimes.* Cambridge, Mass.: Ballinger.

Edwards, Sebastian, and Miguel A. Savastano. 1999. "Exchange Rates in Emerging Economies: What Do We Know? What Do We Need to Know?" Revised version of a paper presented at a Stanford University conference on "Economic Policy Reform: What We Know and What We Need to Know," September 17–19, 1998. Processed.

Krueger, Anne. 1978. *Foreign Trade Regimes and Economic Development: Liberalization Attempts and Consequences.* Cambridge, Mass.: Ballinger.

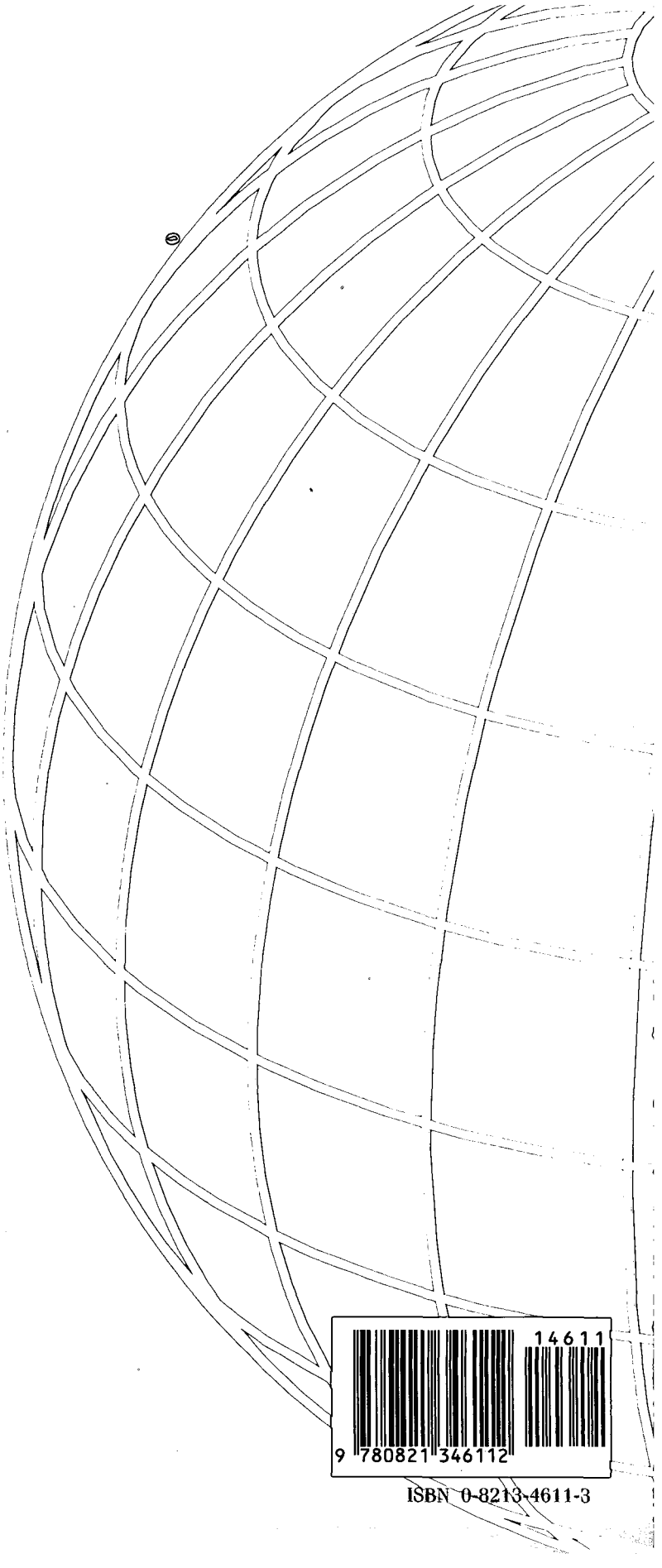
Little, Ian, Tibor Scitovsky, and Maurice Scott. 1970. *The Structure of Protection in Developing Countries: A Comparative Study.* Baltimore, Md.: Johns Hopkins University Press.

Papageorgiou, Demetris, Michael Michaely, and Armeane M. Choksi, eds. 1991. *Liberalizing Foreign Trade.* Oxford, U.K.: Basil Blackwell.

Soros, George. 1998. *The Crisis of Global Capitalism.* New York: Public Affairs.

THE WORLD BANK

1818 H Street N.W.
Washington, DC 20433, U.S.A.
Telephone: (202) 477-1234
Facsimile: (202) 477-6391
Telex: MCI 6415 WORLDBANK
MCI 248423 WORLDBANK
Internet: www.worldbank.org/
E-mail: books@worldbank.org



14611

9 780821 346112

ISBN 0-8213-4611-3